





SOURCE  
ORGANISM  
Taeniopygia guttata  
Taeniopygia guttata  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Archosauria; Aves; Neognathae; Passeriformes; Estrilidae;  
Estrildinae; Taeniopygia.  
REFERENCE  
AUTHORS  
1 (bases 1 to 805)  
Clayton, D.F., Arnold, A.P., Ball, G.F., Brenowitz, E., George, J.M.,  
Mello, C.V., Wade, J., Replegle, K., Lewin, H., Band, M., Hernandez, A.  
and Liu, L.  
The Songbird Neurogenomics Initiative: An Evolving Public Resource  
for Study of Genes, Brain, and Behavior  
Unpublished (2004)  
CONTACT: David F. Clayton  
University of Illinois  
8107 CLSL, 601 S. Goodwin, Urbana, IL 61801, USA  
Tel: 217 244 3668  
Fax: 217 244 1648  
Email: dclayton@uiuc.edu  
Base Calling/Quality Scores: PHRED from Washington University  
Genome Center  
Vector Trimming: Cross\_match from Washington University Genome  
Center PHRAP suite. Low quality bases (Phred score < 20) were  
trimmed from both ends of the sequence by an in-house script.  
This sequence is vector free and at least 200 bp in length. Funded  
by PHS grant # R01 NS045264, 'Songbird Neurogenomics Initiative.'  
PCR Primers  
FORWARD: TAATACGACTCACTATAGG(T7)  
BACKWARD: ATTAACTCTCTAAAG(T3)  
Insert Length: 805 Std Error: 0.00  
Plate: SB02048A2 row: A column: 08  
Seq primer: TAATACGACTCACTATAGG (T7)  
High quality sequence stop: 805.  
FEATURES  
source  
Location/Qualifiers  
1..805  
/organism="Taeniopygia guttata"  
/mol\_type="mRNA"  
/db\_xref="taxon:59729"  
/clone="SB02048A2A08.f1"  
/tissue\_type="brain"  
/dev\_stage="late embryo, post-hatch days 1, 10, 20, 45,  
and adult (pooled)"  
/lab\_host="DH10B"  
/clone\_lib="normalized Keck-Tagu Library SB02"  
/note="Organ: brain; Vector: pBS II SK(+); Site 1:  
EcoRI(5' side of insert); Site 2: NotI (3' side of  
insert); The library was constructed and normalized as  
described by Bonaldo, M.F., Lennon, G. and Soares, M.B.  
(1996), Genome Research 6(9): 791-806. An identifying tag  
was added at the 3' during cDNA synthesis:  
insertAAAAAAAAAAAAAAAAAATGCGA."

ORIGIN  
Query Match 34.4%; Score 542.6; DB 7; Length 805;  
Best Local Similarity 79.4%; Pred. No. 1.2e-107;  
Matches 638; Conservative 0; Mismatches 166; Indels 0; Gaps 0;  
QY 630 GATCATGAGCTGGTATGACGTCATCTGTACGTGGTGTGAGTGGCAACTCATGTA 689  
DB 1 GATCATGAACTGGTATGACCTGCATCTCTATGTGGTGGTGGGAACCTGATGA 60  
QY 690 CAACAGCTTCCCGGGGCTGCCGTGTCGAGAAAGTCTCTGTTCATTTATCGCCACGCCCT 749  
DB 61 CAACAGCTTCCCAACTGCGCGTCTCCAGAAAGTGTGTGTCATCTATGCGACGCACT 120  
QY 750 GTCGTGCTTCGCGCTTCTTAAAGAACTCAAGCGCGTGTCCAAAGTTCAGTCTGCTGTG 809  
DB 121 GTCCTGCTTGTGCGTCTTGAAGAACTCAAGCGAGTCTCAAGTTCAGTTCGCTGTG 180  
QY 810 CACTCTGGCCCACTTCGTATCATATCTCTGTTGATAGCTACTGTTATCGCGGCGG 869  
DB 181 CAGTTAGCCCACTTTGTTCATCAACATCTCTGGTGTGATTCGCTCTCCAGGGCGG 240  
QY 870 CGACTGGGCTGGGAGAGGTCAAGTCTTACATCGACGTCGAAGTTCCTCCATCTCCAT 929

Db 241 TGACTGGGCTGGGACAAAGTCAAGTTTACATGATGTCAAGAAGTTTCCCATCTCCAT 300  
QY 930 TGGCATCATCGTGTTCAGCTACAGCTCTCAGATCTTCTCGCTTCGCTGGAGGGCAATAT 989  
Db 301 TGGCATCATGTTCTTCAGCTACACCTCCAGATCTTTCGCTTCCTTTGGAGGGGAACAT 360  
QY 990 GCAGCAGCCCAAGCAGTTCACATGATGAACTGGACGACATCGCAGCTCGCTGCT 1049  
Db 361 GCAGAACCCCAAGAGTTTCAATGATGAACTGGACTCACATGACAGCTTGCACTCT 420  
QY 1050 CAAGGCTCTTTCGCGCTCGTTCAGCTACCTCAGCTGGGCCGACGAGCAAGAGGTCAT 1109  
Db 421 TAAGGCACTCTTTCGCTTGGTGGCTACCTGACCTGGGCTGATGAGACCAAGGAGTCTAT 480  
QY 1110 CACGGTAACCTGCGCGCTCCATCGCGCGTGGTCAACATCTTCTGTTGGCCCAAGGC 1169  
Db 481 TACAGAACACTTGCCATCCACATAGGGCAGTAGTCAATATTTCTTGGTGGCCAAAGC 540  
QY 1170 GCTGTTGCTTATCTCTGCGCTTCTTTCGCGCTGTCGAGGTGCTGGAGAGTTCGCTCTT 1229  
Db 541 CTGCTCTACACCTTTCGCTTCTTTCGAGCTGTAGAAGTCTCGGAGCGATCCCTTTT 600  
QY 1230 CCAGAAAGCAGCGCGCTTTTCCCGCTGCTTACAGCGGCGAGCGGCGCTGAAGTC 1289  
Db 601 CCAAGATGAAACAGGGCTTTCTTCCCAACTGCTATGCGGGGTGACGCGGCGCTCAAAATC 660  
QY 1290 CTGGGGCTGACGCTGCGCTGCGCGCTGCTTTCACGCTGCTTCATGGCCATTTATGT 1349  
Db 661 CTGGGGACTCACCTTCAGATGTGCGCTGTTTTCACCTGCTTCATGGCTTANNATGT 720  
QY 1350 GCGGCACCTTCGCGTGTCTGCGGCTCACCGGCGCTCACGGGCGCGGCTCTGTTT 1409  
Db 721 CCCCATTTCCTCTGATGGGCTTACTGGGAGCTCACAGCGCNGNNNCTCTGTTT 780  
QY 1410 CTGCTGCCAGCTCTTTTTCACCT 1433  
Db 781 CTGCTGCCAGCTCTTTCACCT 804

RESULT 13  
BU709443 689 bp mRNA linear EST 09-JUL-2003  
UI-M-EVO-cbg-e-06-0-UI.r1 NIH\_BMAP\_EVO Mus musculus cDNA clone  
IMAGE:6809167 5', mRNA sequence.  
BU709443  
ACCESSION BU709443  
VERSION BU709443.1 GI:23642916  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 689)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Dr. James Lin, University of Iowa  
cDNA Library prepared by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
This clone was contributed by the Brain Molecular Anatomy Project  
(BMAP)  
Seq primer: pyx-5.  
Location/Qualifiers  
1..689  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6"

FEATURES  
source  
Location/Qualifiers  
1..689  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6"

```
Db CCTCAAGAACCTCAAGACCGGTGTCCTCAAGTTTAGCTGCTGCACCTAGGCATTCAT 365
QY CATCAATATCTGTCTAGCTACTGTCTATCGGGCGCGGCACTGGGCTGGAGAA 887
Db CATCAACATCTTAGTGTCTGCTACTGCTGTCTCAGGGCGGAGACTGGGCTGGACAA 425
QY GGTCAAGTTTACATCGACGTCAGAGAGTTCCCATCTCCATTCGGCATCATCGTGTTCAG 947
Db GGTCAAGTTTACATCGACGTCAGAGAGTTCCCATCTCCATTCGGCATCATCGTGTTCAG 485
QY CTACACGTTCTCAGATCTTCTGCTCTGCTGAGGCAATATGACAGCCACCGAGATT 1007
Db CTACACCTCGCAGATCTTCTGCTCTGCTGAGGCAATATGACAGCCACCGAGATT 545
QY CCATGATGATGAATGAGCGCATCGACGATCGGAGGCAATATGACAGCCACCGAGATT 1067
Db CCATGATGATGAATGAGCGCATCGACGATCGGAGGCAATATGACAGCCACCGAGATT 605
QY CGTGGCTTACCTACCTCGGCGGAGAGACCAAGGAGTATGACAGCCACCGAGATT 1127
Db GGTGGCTTACCTACCTCGGCGGAGAGACCAAGGAGTATGACAGCCACCGAGATT 665
QY CTCATCGGCGGCTGTGCTCAACATCTTCTGCTGAGGCAATATGACAGCCACCGAGATT 1187
Db CAGCATCGGCGGCTGTGCTCAACATCTTCTGCTGAGGCAATATGACAGCCACCGAGATT 725
QY GCAATCTTCTGCTGAGGCTGTGAGAGAGTGTGCTGCTGAGGCAATATGACAGCCACCGAGATT 1247
Db GCGCTTCTGCTGAGGCTGTGAGAGAGTGTGCTGCTGAGGCAATATGACAGCCACCGAGATT 785
QY CTTTCTTCTGCTGAGGCTGTGAGAGAGTGTGCTGCTGAGGCAATATGACAGCCACCGAGATT 1298
Db GTTCTTCTGCTGAGGCTGTGAGAGAGTGTGCTGCTGAGGCAATATGACAGCCACCGAGATT 836
```

```
RESULT 11
CK619909 624 bp mRNA linear EST 26-JAN-2004
LOCUS mk29g12.y1 Mouse retina, unamplified: mk/ml Mus musculus cDNA clone
DEFINITION mk29g12 5', mRNA sequence.
ACCESSION CK619909
VERSION CK619909.1 GI:41340795
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Ida H., Boylan S., Weigel A., Smit-McBride Z., Chao A., Gao J.,
Buchoff P., Wistow G. and Hjelmeland L.
Expressed sequence tag analysis of mouse retina
Unpublished (2004)
Contact: Wistow G
Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 496 0078
Email: graeme@helix.nih.gov
Plate: 29 row: g column: 12
Seq primer: M13RP1 reverse primer (ABI).
Location/Qualifiers
1. 624
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57Bl6J"
/db_xref="taxon:10090"
/clone="mk29g12"
/sex="Male"
/tissue_type="retina"
/dev_stage="Adult"
/lab_host="EMDH10B"
```

## FEATURES

## source

```
RESULT 12
CK309005
LOCUS SB02048A08.f1
DEFINITION guttata cDNA clone SB02048A08.f1 5, mRNA sequence.
ACCESSION CK309005
VERSION CK309005.1 GI:44818579
KEYWORDS EST.
```

```
/clone_lib="Mouse retina, unamplified: mk/ml"
/note="Organ: Eye; Vector: pSport1; 270ug total RNA was
extracted from 200 adult male mouse retinas. A
directionally cloned cDNA library in the pSport1
vector (Life Technologies) was constructed at Bioserve
Biotechnology (Laurel MD) essentially following the
protocols of the SuperScript Plasmid System full details
of which are contained in the manufacturer's instruction
manual (http://www.lifetech.com/). First strand synthesis
was carried out using a Not I primer-adaptor
[5'-pGACTAGTTCTAGATCGGAGCGGCC(T)15-3']. cDNA was
cloned in Not I/Sal I sites. EST analysis was performed on
the unamplified library at the NIH Intramural Sequencing
Center (NISC)."
```

## ORIGIN

```
Query Match 34.5%; Score 544; DB 7; Length 624;
Best Local Similarity 92.0%; Pred. No. 5 6e-108;
Matches 574; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 947 GCTACAGTCTCAGATCTTCTGCTGCTGAGGCAATATGACAGCCACCGAGATT 1006
Db 1 GCTACAGTCTCAGATCTTCTGCTGCTGAGGCAATATGACAGCCACCGAGATT 60
QY 1007 TCCACTGTCATGATGAATGAGCGCATCGACGCTGCTCAAGGGCTCTTCGCGC 1066
Db 61 TCCACTGTCATGATGAATGAGCGCATCGACGCTGCTCAAGGGCTCTTCGCGC 120
QY 1067 TCGTCTGCTACCTCACTCGGCGGAGAGACCAAGGAGTATGACAGCCACCGAGATT 1126
Db 121 TCGTCTGCTACCTCACTCGGCGGAGAGACCAAGGAGTATGACAGCCACCGAGATT 180
QY 1127 GCTCCATTCGCGCGGCTGTGCTCAACATCTTCTGCTGAGGCAATATGACAGCCACCGAGATT 1186
Db 181 GCTCCATTCGCGCGGCTGTGCTCAACATCTTCTGCTGAGGCAATATGACAGCCACCGAGATT 240
QY 1187 TGCATTTCTTCTGCTGAGGCTGTGAGAGTGTGCTTCCAGGAGGAGCGCGC 1246
Db 241 TGCCTTTCTTCTGCGGCGCTGAGAGTGTGAGAGTGTCTTCCAGGAGGAGCGCGC 300
QY 1247 CTTTCTTCTGCGGCGCTGTACAGCGGCGGAGAGTGTGAGAGTGTCTTCCAGGAGGAGCGCGC 1306
Db 301 CTTTCTTCTGCGGCGCTGTATGAGGCGGAGCGGTGCTTAAAGTCTTGGGGCTGACGCTGC 360
QY 1307 GCTGCGGCTGTGCTGTCTTCAAGCTGTATGAGGCAATATGAGGCGCACTTCGCGCTGC 1366
Db 361 GCTGCGGCTGTGCTGTCTTCAAGCTGTATGAGGCAATATGAGGCGCACTTCGCGCTGC 420
QY 1367 TCATGGGCTGTACCGGCGGAGGCTGTGAGGCGGCGGCTGTGCTTCTTCTGCTGCGGAGGCTCT 1426
Db 421 TCATGGGCTGTACCGGCGGAGGCTGTGAGGCGGCGGCTGTGCTTCTTCTGCTGCGGAGGCTCT 480
QY 1427 TCCACTGCGGCTGTCTGCGGCAAGCTGTGAGGCAATATGAGGCGCACTTCTGAGCTGCGCA 1486
Db 481 TCCACTTGGGCTGTCTTCTGCGGCAAGCTGTGAGGCAATATGAGGCGCACTTCTGAGCTGCGCA 540
QY 1487 TCTTCTGTCATCGGCGGCAATGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTCA 1546
Db 541 TCTTCTGTCATCGGCGGCAATGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTCA 600
QY 1547 TCGAGGCTTACCGAACCAGCGG 1570
Db 601 TCGAGGCTTACCGAACCAGCGAG 624
```

```
CK309005 805 bp mRNA linear EST 01-MAR-2004
SB02048A08.f1 normalized Keck-Tagu Library SB02 Taeniolypgia
guttata cDNA clone SB02048A08.f1 5, mRNA sequence.
CK309005
CK309005.1 GI:44818579
EST.
```

before RNA extraction and purification (Tri-reagent method). cDNA was primed using oligo-dT primer: 5'-pGACTAGTCTTAGATCGGAGCGCGCC(T)25-3' and cloned into the EcoRV/NotI sites of pExpress-1. Size-selection >1.25kb resulted in an average insert size of 2.18 kb. This primary library is not normalized (normalized library is NIH MGC 255) and was constructed by Express Genomics (Frederick, MD). Note: this is a NIH\_MGC library"

## ORIGIN

Query Match 37.4%; Score 590; DB 7; Length 763;  
Best Local Similarity 91.1%; Pred. No. 5.4e-118;  
Matches 638; Conservative 0; Mismatches 60; Indels 2; Gaps 1;

QY 702 GGGGTCGCCGTCGCGAGAGTCTGTGTCATTATCCGACGGCGGCTGCTGCTGCTG 761  
DB 1 GGGGTCGCCGTCGCGAGAGTCTGTGTCATTATCCGACGGCGGCTGCTGCTGCTG 60

QY 762 CGCCTTCTCTTAAGAACCTCAAGCGCGTGTCCAAAGTTTCAGTCTGTGTGCACTCTGCCCCA 821  
DB 61 CGCCTTCTCTTAAGAACCTCAAGCGCGTGTCCAAAGTTTCAGTCTGTGTGCACTCTGCCCCA 120

QY 822 CTTCTGTCATCAATATCTCTGGTTCATAGCCCTACTGTCTATCGCGGCGCGAGCTGGGCTG 881  
DB 121 CTTCTGTCATCAATATCTCTGGTTCATAGCCCTACTGTCTATCGCGGCGCGAGCTGGGCTG 180

QY 882 GAGAGAGTCAAGTTCTACATCGAGTCAAGAGTTCCCAATCTCCATTCGATCGATCATGCT 941  
DB 181 GAGAGAGTCAAGTTCTACATCGAGTCAAGAGTTTCCTATCTCCATTCGATCGATCATGCT 240

QY 942 GTTTCAGTACAGCTCTCAGATCTTCTGCTTCTGCTGAGGCAATATGACAGCCCGAG 1001  
DB 241 GTTTCAGTACAGCTCTCAGATCTTCTGCTTCTGCTGAGGCAATATGACAGCCCGAG 300

QY 1002 CGAGTTCCATCGATGATGAATCTGACGACATCGAGCGCTGCTGCTCAAGGGCTCTTT 1061  
DB 301 CGAATTCATCGATGATGAATCTGACGACATCGAGCGCTGCTGCTCAAGGGCTCTTT 360

QY 1062 CGGCTGCTGCTACCTACCTGCGGCGAGACGAGACCGAGGAGTCAACCGATAACT 1121  
DB 361 CGGCTGCTGCTACCTACCTGCGGCGAGACGAGACCGAGGAGTCAACCGATAACT 420

QY 1122 GCGCGCTCCATCCGCGCGTGTCAACATCTTCTGCTGCGCAAGCGCTGCTGCTCA 1181  
DB 421 GCGCGCTCCATCCGCGCGTGTCAACATCTTCTGCTGCGCAAGCGCTGCTGCTCA 480

QY 1182 TCCTCTGCCATCTTTTCGCGCTGTGAGGTGCTGGAGAGTCTGCTTCCAGAGGCGAG 1241  
DB 481 CCGGTTGCGCTTCTTCGCGCGCTGCAAGTGTGGAGAGTCTCTCTCCAGAGGCGAG 540

QY 1242 CCGCGCTTTTTCGCGCGCTGTACAGCGGCGAGCGGCGCTGAGTCTGCGGCTGAC 1301  
DB 541 TCGTGCCTTTCTTCCCGCGCTGTACGCGGCGAGCGGCTGAGTCTTTCGCGGCTGAC 600

QY 1302 GCTGCGCTGCGGCTGCTGCTTTCAGCTGTCTATGCGCATTTATGTCGCGCACTTCG 1361  
DB 601 GCTGCGCTGCGGCTGCTGCTTTCAGCTGTCTATGCGCATTTATGTCGCGCACTTCG 660

QY 1362 GTGTGCTATGGGCTTCAACGCGAGCCTCAAGCGCGCGCGC 1401  
DB 661 GCTGCTCATGG--CCTCACGGGCACTTCAAGCGGCGCGC 698

## RESULT 10

CN503909  
LOCUS 22405600 NIH\_ZGC\_9 Danio rerio cDNA clone IMAGE:7273715  
DEFINITION 5', mRNA linear EST 28-APR-2004  
5', mRNA sequence.  
CN503909  
ACCESSION CN503909.1 GI:46816533  
VERSION CN503909.1  
KEYWORDS EST.  
SOURCE Danio rerio (zebrafish)  
ORGANISM Danio rerio

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.  
1 (bases 1 to 937)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: Chi-Bin Chien  
cDNA Library Preparation: Dr. Sumio Sugano  
cDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
DNA distribution: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
plate: LLAM15261 row: i column: 09  
High quality sequence start: 8  
High quality sequence stop: 698.

FEATURES  
source

1..937  
/organism="Danio rerio"  
/mol\_type="mRNA"  
/db\_xref="taxon:7955"  
/clone="IMAGE:7273715"  
/tissue\_type="neural retina, retinal pigment epithelium, lens and overlying skin, pooled embryos"  
/lab\_host="DH10B Tona"  
/clone\_lib="NIH\_ZGC\_9"  
/note="Organ: eye; Vector: pME18S-PL3; Site 1: DraIII; Site 2: DraIII; 1st strand cDNA was primed with an oligo(dT) primer  
[GCGCTGAGACGGCTATGCGCTTTTCTTTTCTTTT];  
double-stranded cDNA was ligated to a DraIII adaptor [GGCUCACUGG], digested and directionally cloned into distinct DraIII sites of the pME18S-PL3. Library was size selected for 1.0 kb, with an average insert size of ~1.2kb, and is not amplified. Library constructed by Yutaka Suzuki (University of Tokyo Institute of Medical Science). Custom primers recommended for sequencing: 5' end primer 5'-GGATGTTGCTTTACTTCTA-3' and 3' end primer 5'-CGACTGCGAGTCGAGACA-3'. Note: This is a Zebrafish Gene Collection (ZGC) library"

## ORIGIN

Query Match 36.9%; Score 582; DB 7; Length 937;  
Best Local Similarity 81.2%; Pred. No. 3.1e-116;  
Matches 675; Conservative 0; Mismatches 156; Indels 0; Gaps 0;

QY 468 CTTTGTGCTGCTACACGGCAAGATCTCTATCGGTCCTGTACGAGGAGAGTGAAGA 527  
DB 6 CTTTGTGCTGCTACACGGCAAGATCTCTATCGCTCTTATGAGAGAGAGAGAGA 65

QY 528 CGGCGAGTGTGCGCGTGGGAGTCTGTACGTGCGCATAGCAACGCTGCTGCGCCCC 587  
DB 66 CGGCGAGTGTGCGGAGTGAAGATCTGTACGTGAGACATCGCAACGCTGCTGCGCTC 125

QY 588 GCGCTTCCCAACGCTGCGGCGCGAGTGTGAGAACGTAGCGCAGATCATCAGCTGTGTAT 647  
DB 126 GCGCTTCCCGGCTCGGAGGACACGTCTGCAACGTGGCACAGATCATCAGCTGTGTAT 185

QY 648 GACGTGCTATCTGTAGTGTGCTGAGTGGCAACCTCATGTACACAGCTTCCCGGGCT 707  
DB 186 GACGTGCTATTTGTAGTGTGCTGAGTGGCAACCTCATGTACACAGCTTCCCGGGCT 245

QY 708 GCGCGTGTGCGAGAGTCTGCTGCTCAATTATCCGACGGCGCTGCTGCTGCTGCGCTT 767  
DB 246 CCGAGTGTCCGAGAGTCTGCGGCGCATCATCGCCACCGCGCTCTCTGCTGCGCTT 305

QY 768 CCTTAAGAACCTCAAGGCGCGTGTCCAAGTTCAGTCTGTGTGACTCTGCGCCACTTCGT 827

KEYWORDS  
SOURCE  
ORGANISM

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

EST.  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 675)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-r@mail.nih.gov  
Tissue Procurement: Dr. James Lin University of Iowa  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone distribution: Distribution information can be found at  
<http://genome.uiowa.edu/distribution/mousefl.html>  
This clone was contributed by the Brain Molecular Anatomy Project  
(BMAP)  
Seq primer: pYX-5.  
Location/Qualifiers  
1..675  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6"  
/db\_xref="taxon:10090"  
/clone="IMAGE:30681555"  
/tissue\_type="whole eye"  
/dev\_stage="newborn (1, 5, 15 days)"  
/lab\_host="DH10B (T1 phage resistant)"  
/clone\_lib="NIH\_BMAP\_HW0"  
/notes="Organ: Eye; Vector: pYX-Asc; Site 1: EcoR I;  
Site 2: Not I; The library was constructed according  
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
1996. Denatured RNA was size fractionated on a 1% agarose  
gel. First strand cDNA synthesis was primed with oligo-dT  
primer containing a Not I site. Double strand cDNA was  
size selected according to mRNA size fraction, ligated  
with EcoR I adaptor, digested with NotI and then cloned  
directionally into pYX-Asc vector. The library tag  
sequence located between the Not I site and the polyA tail  
is AATAATTACG. This library was created for the University  
of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the  
Developing Mouse Nervous System', supported by National  
Institute of Mental Health (NIMH)."  
37.6%; Score 593.4; DB 7; Length 675;  
Best Local Similarity 92.4%; Pred. No. 9.7e-119;  
Matches 624; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

ORIGIN

Query Match  
Best Local Similarity 92.4%; Pred. No. 9.7e-119;  
Matches 624; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 902 TCGACGTCAAGAGTTCCCATCTCCATTTGGCATCATCGTTTACGCTACAGTCTCAGA 961  
DB 1 TCGACGTCAAGAGTTCCCATCTCCATTTGGCATCATCGTTTACGCTACAGTCTCAGA 60  
QY 962 TCTTCTGCTCTCGCTGGAGGGAATATGACAGAGCCAGCGAGTTCCAGTCCATCATGA 1021  
DB 61 TCTTCTGCTCTCTCGAAGGGAATATGACAGAGCCAGCGAGTTCCAGTCCATCATGA 120  
QY 1022 ACTGGACGCACATCGCAGCTCGTGTCTCAAGGGCTCTTTCGGGCTCGTCCGCTACCTCA 1081  
DB 121 ACTGGACGCACATCGCAGCTCGTGTCTCAAGGGCTCTTTCGGGCTCGTCCGCTACCTCA 180  
QY 1082 CTTGGGCGCAGCAGACCAAGAGGTATATCAGGATAACTCTGCCGGCTCATCCGGCGCG 1141  
DB 181 CTTGGGCGCAGCAGACCAAGAGGTATATCAGGATAACTCTGCCGGCTCATCCGGCGCG 240  
QY 1142 TGGTCAACATCTTTCTGGTGGCCAGCGCTGTGTCCTATCTCTGCTGCTTTTGGCG 1201  
DB 241 TGGTCAACATCTTTCTGGTGGCCAGCGCTGTGTCCTATCTCTGCTGCTTTTGGCG 300  
QY 1202 CTGTCGAGGTGCTGGAGAGTCTGCTCTTCAGGAAGCGAGCGCGCTTTTCCCGGCT 1261

DB 301 CCCTCGAAGTGTGGAGAAAGTCTCTCTCCAGGAAGCGAGTCCGGCGCTTCTTCCCGCCT 360  
QY 1262 GCTACAGCGCGCAGCGCGCTCTGAAGTCTCTGGGGGCTGAGCGTGGCGCTCGCTCGTCG 1321  
DB 361 GCTATGAGCGCAGCGTTCGCTTAAGTCTCTGGGGGCTGAGCGTGGCGCTCGCTGGTGG 420  
QY 1322 TCTTACAGCTGCTCATGGCCATTTATGTGCGGCACCTTTCGGGCTGCTCATGGGCTCACCG 1381  
DB 421 TCTTACAGCTGCTCATGGCCATTTATGTGCGGCACCTTTCGGGCTGCTCATGGGCTCACCG 480  
QY 1382 GAGCTCTCAGCGCGCGCGCTCTGTCTTCTGTCGCCAGCTCTTTCACCTCGGCTGTC 1441  
DB 481 GAGCTCTCAGCGCGCGCGCTCTGTCTTCTGTCGCCAGCTCTTTCACCTCGGCTGTC 540  
QY 1442 TCTGGCGCAGCTGCTGTCGCCAGCTCTTCTTCGACGTCGCGCATCTTCTGTCATGGCG 1501  
DB 541 TCTGGCGCAGCTGCTGTCGCCAGCTCTTCTTCGATGTGGCCATCTTCTGTCATGGCG 600  
QY 1502 GCATCTGCGAGCTGTCCGGCTTCGTGCACCTCCCTCGAGGGCTCATCGAGGCTACCGAA 1561  
DB 601 GCATCTGCGAGCTGTCCGGCTTCGTGCATTCATCTCGAGGCTCATCGAGGCTACCGAA 660  
QY 1562 CCAACGCGAGGACT 1576  
DB 661 CCAACGCGAGGACT 675

RESULT 9  
CO393521  
LOCUS  
DEFINITION  
IMAGE:7314034 5', mRNA sequence.  
CO393521  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Rattus norvegicus (Norway rat)  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 763)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: cgabbs-r@mail.nih.gov  
Tissue Procurement: Drs. Josef Lazar & Howard Jacob, Medical  
College of Wisconsin  
cDNA Library Preparation: Express Genomics  
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM15366 row: a column: 08  
High quality sequence stop: 738.  
Location/Qualifiers  
1..763  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10116"  
/clone="IMAGE:7314034"  
/sex="both"  
/tissue\_type="Brain - Pooled from several tissues from one  
or more individuals"  
/lab\_host="DH10B TonA"  
/clone\_lib="NIH\_MGC\_254"  
/note="Organ: Brain/CNS; Vector: pExpress-1; Site 1:  
EcoRV; Site 2: NotI; RNA obtained from brain tissue of 8  
wk old animal. Tissues were snap-frozen and kept at -80C







Qy	626	CGCAGATCATCGAGCTGGTGTATGACGTGCATCTCTGTA	CGTGGTGGTGTGAGTGCAACCTCA	685
Db	480	CGCAGATCATCGAGCTGGTGTATGACGTGCATCTCTGTA	CGTGGTGGTGTGAGTGCAACCTCA	539
Qy	686	TGTACACAGCTTCCCGGGGCTGCCCGTGCAGAACTCTGTGTC	CAATTACGCCACGG	745
Db	540	TGTACACAGCTTCCCGGGGCTGCCCGTGCAGAACTCTGTGTC	CAATTACGCCACGG	599
Qy	746	CCGTGCTGCTGCTTCCCTTGCGCCTTCTTAAAGAACTCAAG	CGCGCTGTCCAAAGTTCAGTCTGC	805
Db	600	CCGTGCTGCTGCTTCCCTTGCGCCTTCTTAAAGAACTCAAG	CGCGCTGTCCAAAGTTCAGTCTGC	659
Qy	806	TGTGCACCTCGCCCACTTCGTGCATCAATATCCTGGTCATA	846	
Db	660	TGTGCACCTCGCCCACTTCGTGCATCAATATCCTGGTCATA	700	

RESULT 4	848 bp	linear	EST 11-JUN-2003
CD557601	CP557601	mrNA	
LOCUS	AGENCOURT 14413666	NIH MGC 180 Homo sapiens	cDNA clone
DEFINITION	IMAGE:30387417	5' . mrNA sequence.	

ORGANISM	HOMO SAPIENS
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ 1 (bases 1 to 848)
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Daniela S. Gerhard, Ph.D. Office of Cancer Genomics National Cancer Institute / NIH Bldg. 31 Rm10A07 Bethesda, MD 20892 Email: cgapbs-r@mail.nih.gov Tissue Procurement: Dr. Michael Brownstein cDNA Library Preparation: Invitrogen Corp cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: NDAM457 row: k column: 10 High quality sequence stop: 569.

ORIGIN	Query Match	41.7%	Score 658.2;	DB 6;	Length 848;
	Best Local Similarity	99.5%	Pred. No. 7.5e-133;		
	Matches 660;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;
Qy	916	TTCCCCATCTCATTGGCATCATCGTGTTCAGTACACGTCCTCAGATCTTCCTCGCTTCG	975		
Db	1	TTCCCCATCTCATTGGCATCATCGTGTTCAGTACACGTCCTCAGATCTTCCTCGCTTCG	60		
Ov	976	CTGGAGGGCAATATGCACAGCCAGCGAGTTCCACTGCATGATGAATCGACGCACATC	1035		

Db	61	1036	121	1096	181	1156	241	1216	301	1276	361	1336	421	1396	481	1456	541	1516	601	1576	661	
Db	CTGAGGGCAATATGACGAGCCAGCGAGTTCCATGATGATGAACGTGAGCGCACATC	GCAGCCTGCGTGTCTCAAGGGCCCTTTCGGCGCTCGTGGCCCTACCTCACTGGGCGGACGAG	GCAGCCTGCGTGTCTCAAGGGCCCTTTCGGCGCTCGTGGCCCTACCTCACTGGGCGGACGAG	ACCAAGAGGTCTATCAGGATAACCTGCCGGGTCCATCGCGGCCGTGTGTCACATCTTTT	ACCAAGAGGTCTATCAGGATAACCTGCCGGGTCCATCGCGGCCGTGTGTCACATCTTTT	CTGGTGCCAAAGGGCTGTGTCCTATCTCTGCCATCTTTTGGCCGCTGTCGAGGTGCTG	CTGGTGCCAAAGGGCTGTGTCCTATCTCTGCCATCTTTTGGCCGCTGTCGAGGTGCTG	GAGAAGTCGCTCTTCCAGGAAGGACGCGCGCCCTTTTTCGGGCCCTGTACAGCGGCGAC	GAGAAGTCGCTCTTCCAGGAAGGACGCGCGCCCTTTTTCGGGCCCTGTACAGCGGCGAC	GAGAGTCGCTCTTCCAGGAAGGACGCGCGCCCTTTTTCGGGCCCTGTACAGCGGCGAC	GGGCGCTGAAGTCCTGGGGGCTGACGCTGCGTGCGGCTCGTCTGTCACGCTGCTC	GGGCGCTGAAGTCCTGGGGGCTGACGCTGCGTGCGGCTCGTCTGTCACGCTGCTC	ATGGCCATTATGTGCGGCACTTTCGGCGCTGCTCATGCGGCTCACCGGACGCTCACGGGC	ATGGCCATTATGTGCGGCACTTTCGGCGCTGCTCATGCGGCTCACCGGACGCTCACGGGC	GCGGCGCTGTTTCTTGCTGCGCCAGCGCTCTTTTACGTCGCGCTGCTCTGGCGCAAGCTG	GCGGCGCTGTTTCTTGCTGCGCCAGCGCTCTTTTACGTCGCGCTGCTCTGGCGCAAGCTG	CTGTGGCACCAGTCTTCTTTCGACGTGCGCCATCTTCTGTCATCGCGGCGCATCTGCAGCGTG	CTGTGGCACCAGTCTTCTTTCGACGTGCGCCATCTTCTGTCATCGCGGCGCATCTGCAGCGTG	TCGCGCTTCGTGCACTCCCTCGAGGGGCTCATCGAAGCCTTACCGAACACCGCGGAGGAC	TCGCGCTTCGTGCACTCCCTCGAGGGGCTCATCGAAGCCTTACCGAACACCGCGGAGGAC	TAG 1578	TAG 663

RESULT 5

B0573293

LOCUS

DEFINITION

778 bp

linear

EST 19-JUN-2002

UI-M-PD0-byg-h-01-0-UI.r1 NIH BMAP\_F00 Mus musculus cDNA clone

IMAGE:5717808 5', mRNA sequence

RESULT 5  
 BQ573293  
 LOCUS  
 DEFINITION  
 BQ573293 778 bp mRNA linear EST 19-JUN-2002  
 UI-M-ED0-byg-h-01-0-UI.r1 NIH-BMAP\_FD0 Mus musculus CDNA clone  
 IMAGE:5717808 5', mRNA sequence.  
 BQ573293  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 BQ573293.1 GI:21476610  
 EST.  
 Mus musculus (house mouse)  
 ORGANISM  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 778)  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT  
 NIH-MGC <http://mgc.nci.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: Dr. James Lin, University of Iowa  
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 This clone was contributed by the Brain Molecular Anatomy Project  
 (ENAP)  
 seq primer: pyX-5.  
 Location/Qualifiers  
 FEATURES



/clone\_lib="NIH\_MGC\_94"  
/lab\_host="DH10B"  
/note="Vector: PCMV-SPORT6"

## ORIGIN

```
Query Match      53.8%; Score 849.2; DB 3; Length 1730;
Best Local Similarity 92.0%; Pred. No. 1.9e-174;
Matches 896; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

QY 605 GCGCGCGAGTGTGAACGAGTAGCGGCGAGATCATCGAGCTGTGTATGACGTGATCCTGTACG 664
Db 1730 GCGCGCGCGTGTCAATGTGGCGCGAGATCATCGAGCTGTGTATGACGTGATCTTGTACG 1671

QY 665 TGGTGTGAGTGCACACCTCATGTACACAGCTTCCCGGGCTGCCGCTGTCGACAGAGT 724
Db 1670 TGTGTGTAGCGCGCAACCTCATGTACACAGTTCCTCGGGGCTGCCGCTGTCGACAGAGT 1611

QY 725 CTGTGTTCATTATCGCACGCGCTGCTGCTGCTTCCCTTGCCTTCCCTTAAGAACCTCAAGG 784
Db 1610 CTGTGTTCATTATCGCACAGCGGTGCTGCTGCTGCTTCCCTTGCCTGAGATCTCAAGG 1551

QY 785 CCGTGTCCAAGTTCAGTCTGTGTGACATCTGCGCCACTTCTGTCATFCAATATCCTGTGCA 844
Db 1550 CCGTGTCCAAGTTCAGTCTGTGTGACATCTGTCGTGCGCCACTTCTGTCATCAACATCTGTGCA 1491

QY 845 TAGCCTACTGTATCGCGGGCGCGGACTGCGCTGGGAGAGTCAAGTTCATCATCG 904
Db 1490 TCGCTTACTGTCTCTCGCGCGCGGTGATTTGGGCTGGGAGAGTGAAGTTCATCATCG 1431

QY 905 AGTCAAGAAGTTCCTCATCTCCATTGGCATCATGTTTACGTACAGCTCTCAGATCT 964
Db 1430 AGTCAAGAAGTTCCTCATCTCCATTGGCATCATGTTTACGTACAGCTCGAGATCT 1371

QY 965 TCTGTCTTCTGTGAGGCAATATGACAGCCAGCGAGTTCCACTGCATGATGAAT 1024
Db 1370 TCTGTCTTCTGTGAGGCAATATGACAGCCAGCGAGTTCCACTGCATGATGAAT 1311

QY 1025 GACGACATCGCAGCTGCTGCTCAAGGGCTCTTCGGCTCGTCCGCTACCTCACT 1084
Db 1310 GACACACATCGCGGCTGCTGCTCAAGGGCTCTTCGGCTCGTCCGCTACCTCACT 1251

QY 1085 GCGCGACGAGCAAGGAGGTATCATCAGGTAACCTGCCGCTCATCCGCGCTGCGCGCTGG 1144
Db 1250 GCGCGACGAGCAAGGAGGTATCATCAGGTAACCTGCCGCTCATCCGCGCTGCGCGCTGG 1191

QY 1145 TCAACATCTTCTGTGTGGCAAGGCGCTGTGTCTTATCTCTGCTGCTATCTTTTGGCGTG 1204
Db 1190 TCAACCTCTTCTGTGTGGCAAGGCGCTGTGTCTTATCTCTGCTGCTATCTTTTGGCGTG 1131

QY 1205 TCGAGTGTGTGAGAGTGTCTTCCAGAGAGGAGCGCGCTTTTCCGCGCTGCT 1264
Db 1130 TCGAGTGTGTGAGAGTGTCTTCTTCCAGAGAGGAGTGTGCGCTTTTCCGCGCTGCT 1071

QY 1265 ACAGCGGCGACGCGCGCTGAGTCTGTGGGCTGACGTGCGCTGCGGCTCGTCT 1324
Db 1070 ATGGAGGCGACGCTGCTTAAAGTCTGTGGGCTGACGTGCGCTGCGGCTGCTGCT 1011

QY 1325 TCAAGTGTCTATGGCCATTTATGTCCGCACTTTCGCTGTCTCATGAGGCTCACCGCA 1384
Db 1010 TCAAGTGTCTATGGCCATTTATGTCCGCACTTTCGCTGTCTCATGAGGCTCACCGCA 951

QY 1385 GCTTACGCGGCGCGGCTGCTGTTTCTGTGCGCAGGCTCTTTTACCTGCGGCTGCT 1444
Db 950 GCTTACGCGGCGCGGCTGCTGTTTCTGTGCGCAGGCTCTTTTACCTGCGGCTGCT 891

QY 1445 GCGCAAGTGTGTGTGCAACAGTCTTCTTGAAGTGTGCTTCTTGTGCTGCGGCA 1504
Db 890 GCGCAAGTGTGTGTGCAACAGTCTTCTTGAAGTGTGCTTCTTGTGCTGCGGCA 831

QY 1505 TCTGCAAGCTGTCCGCTTCTGTGCACTTCCCTTCAGGCGCTTCTTGAAGCTTACCGAACCA 1564
Db 830 TCTGCAAGCTGTCCGCTTCTGTGCACTTCTTGAAGCTTCTTGAAGCTTACCGAACCA 771
```

QY 1565 ACGCGAGGACTAG 1578  
Db 770 ACGCAGGACTAG 757

RESULT 2  
CNS0292M

LOCUS  
DEFINITION  
Tetraodon nigroviridis genome survey sequence T7 end of clone  
245P02 of library G from Tetraodon nigroviridis, genomic survey  
sequence.

## ACCESSION

VERSION  
ALI86727.1 GI:7824831

## KEYWORDS

GSS; genome survey sequence.

## SOURCE

Tetraodon nigroviridis

## ORGANISM

Tetraodon nigroviridis  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
Tetraodontoidea; Tetraodontidae; Tetraodon.

## REFERENCE

## AUTHORS

Roest Crolius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C.,  
Bernot, A., Fzanes, C., Wincker, P., Brottier, P., Quetier, F.,  
Saurin, W. and Weissenbach, J.

## TITLE

Estimate of human gene number provided by genome-wide analysis  
using Tetraodon nigroviridis DNA sequence

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

Roest Crolius, H., Jaillon, O., Dasilva, C., Ozouf-Costaz, C.,  
Fzanes, C., Fisher, C., Bouneau, L., Billault, A., Quetier, F.,  
Saurin, W., Bernot, A. and Weissenbach, J.  
Characterization and repeat analysis of the compact genome of the  
freshwater pufferfish Tetraodon nigroviridis

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 3, 2004, 18:33:25 ; Search time 4611 Seconds  
(without alignments)  
12470.591 Million cell updates/sec

Title: US-09-940-919-1  
Perfect score: 1578  
Sequence: 1 atggccacctgtctgcgacg.....gaaccaacgaggactag 1578

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*

1: gb\_est1.\*  
2: gb\_est2.\*  
3: gb\_est3.\*  
4: gb\_est4.\*  
5: gb\_est5.\*  
6: gb\_est6.\*  
7: gb\_est7.\*  
8: gb\_est8.\*  
9: gb\_est9.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	849.2	53.8	1730	BC025013	Mus muscu
2	689.2	43.7	1027	CNS0292M	AL188727 Tetraodon
3	677.8	43.0	700	BG700073	BG700073 602681089
4	658.2	41.7	848	CD557601	CD557601 AGENCOURT
5	638.2	40.4	778	BQ573293	BQ573293 UI-M-PD0-
6	607.8	38.5	658	CD635843	CD635843 56000431J
7	597.6	37.9	1031	CNS03DEA	AL238987 Tetraodon
8	593.4	37.6	675	CO429271	CO429271 UI-M-HW0-
9	590	37.4	763	CO393521	CO393521 AGENCOURT
10	582	36.9	937	CNS03909	CNS03909 AGENCOURT
11	544	34.5	624	CK619909	CK619909 mk29g12.Y
12	542.6	34.4	805	7 CK309005	CK309005 SB02048A2
13	537.4	34.1	689	5 BU709443	BU709443 UI-M-EV0-
14	525.2	33.3	605	6 CB581479	CB581479 AMGNNUC:N
15	520.2	33.3	689	9 AG576538	AG576538 Mus muscu
16	520.2	33.0	1101	9 CNS052RP	AL318526 Tetraodon
17	511.6	32.4	1080	9 CNS05L98	AL342485 Tetraodon
18	509	32.3	767	4 BJ735127	BJ735127 BJ735127
19	503.8	31.9	579	2 AW742427	AW742427 up57a04.Y
20	495.4	31.4	595	8 AZ098746	AZ098746 RPCI-23-1
21	495	31.4	670	4 B1822066	B1822066 603039933
22	464.8	29.5	886	7 CN833492	CN833492 AGENCOURT
23	464.4	29.4	549	8 AQ933231	AQ933231 RPCI-23-2
24	457.2	29.0	693	9 CNS04J95	AL293234 Tetraodon

25	456.6	28.9	863	7	CN838148	CN838148 AGENCOURT
26	444	28.1	508	2	BE981082	BE981082 UI-M-CG0-
27	435.6	27.6	522	8	AZ002297	AZ002297 RPCI-23-3
28	433.4	27.5	751	4	BI488536	BI488536 603021049
29	433	27.4	509	1	AI839048	AI839048 UI-M-AKO-
30	429	27.2	608	4	BI393349	BI393349 p9p1n.pk0
31	419.2	26.6	480	2	BE982265	BE982265 UI-M-CG0p
32	410.2	26.0	471	2	AW120774	AW120774 UI-M-BH2-
33	395	25.0	623	1	AL588280	AL588280 AL588280
34	393	24.9	487	1	AI836076	AI836076 UI-M-AU0-
35	386.2	24.5	447	1	AI834859	AI834859 UI-M-AN1-
36	382.2	24.2	734	5	BU339660	BU339660 603406545
37	381.8	24.2	449	2	AW048723	AW048723 UI-M-BH1-
38	367.2	23.3	867	7	CN833360	CN833360 AGENCOURT
39	365.2	23.1	455	2	BE950161	BE950161 UI-M-CE0-
40	362.6	23.0	456	1	AI837592	AI837592 UI-M-AKO-
41	341.4	21.6	645	5	BU053668	BU053668 UI-M-FD0-
42	337.4	21.4	457	7	H23502	H23502 Ym57h09.r1
43	334	21.2	942	6	CD361004	CD361004 AGENCOURT
44	333	21.1	392	1	AI549592	AI549592 UI-R-C3-t
45	328	20.8	403	2	BF543591	BF543591 UI-R-C3-t

#### ALIGNMENTS

RESULT 1  
BC025013/c  
LOCUS BC025013 1730 bp mRNA linear HTC 07-AUG-2002  
DEFINITION Mus musculus, clone IMAGE:5357287, mRNA.  
ACCESSION BC025013  
VERSION BC025013.1 GI:19035721  
KEYWORDS HTC.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 1730)  
AUTHORS Strausberg,R.  
TITLE Direct Submission  
JOURNAL Submitted (01-MAR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
COMMENT Contact: MGC help desk  
Email: [cgapsb@mail.nih.gov](mailto:cgapsb@mail.nih.gov)  
Tissue Procurement: The Cepko Laboratory  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: [ang@bcm.tmc.edu](mailto:ang@bcm.tmc.edu)  
Gunnarathne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M., Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAC Plate: 54 Row: f Column: 19  
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis  
This clone has the following problem: no polyA-tail.  
Location/Qualifiers  
1. 1730  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5357287"  
/tissue\_type="Eye, retina, mouse strain C57Bl/6"



Db	1372	TGCTACAGCGCGACGCGCGCTGAAGTCCTGGGGGCTGACGCTGCGCTGCGGCTCGTC	1431
Qy	1321	GTCTTACGCTGCTCATGGCCATTATGTGCGCACTTCGCGCTGCTCATGGGCTCACC	1380
Db	1432	GTCTTACGCTGCTCATGGCCATTATGTGCGCACCTCGGCTGCTCATGGGCTCACC	1491
Qy	1381	GGCAGCCTCAGCGGCGCGGCTCTGTTTCTTGTGCCCAGCCTCTTTTCACCTGCGGCTG	1440
Db	1492	GGCAGCCTCAGCGGCGCGGCTCTGTTTCTTGTGCCCAGCCTCTTTTCACCTGCGGCTG	1551
Qy	1441	CTCTGGCGCAAGCTGCTGTGGCACCAAGTCTTTCGACGTGCGCATCTTTCGTCATCGGC	1500
Db	1552	CTCTGGCGCAAGCTGCTGTGGCACCAAGTCTTTCGACGTGCGCATCTTTCGTCATCGGC	1611
Qy	1501	GGCATCTGCAGCGTGTCCGGCTTGTGCACTCCCTCGAGGGCCTCATCGAAGCCTACCGA	1560
Db	1612	GGCATCTGCAGCGTGTCCGGCTTGTGCACTCCCTCGAGGGCCTCATCGAAGCCTACCGA	1671
Qy	1561	ACCAACGCGGAGGACTAG	1578
Db	1672	ACCAACGCGGAGGACTAG	1689

Search completed: November 3, 2004, 22:38:58  
Job time : 788 secs

APPLICANT: Edinger, Shlomit  
APPLICANT: Gerlach, Valerie  
APPLICANT: Ellerman, Karen  
APPLICANT: MacDougall, John R  
APPLICANT: Gunther, Erik  
APPLICANT: Millet, Isabelle  
APPLICANT: Stone, David J  
APPLICANT: Smithson, Glennda  
APPLICANT: Szekeres Jr, Edward S  
APPLICANT: Ji, Weizhen  
TITLE OF INVENTION: No. US20040068095A1el Human Proteins, Polynucleotides Encoding Th  
TITLE OF INVENTION: Methods Of Using The Same  
FILE REFERENCE: 21402-141 CIP  
CURRENT APPLICATION NUMBER: US/10/096,625  
PENDING FILING DATE: 2002-03-13  
PRIORITY APPLICATION NUMBER: 09/972,211  
PRIORITY FILING DATE: 2001-10-05  
PRIORITY APPLICATION NUMBER: 60/238,325  
PRIORITY FILING DATE: 2000-10-05  
PRIORITY APPLICATION NUMBER: 60/238,323  
PRIORITY FILING DATE: 2000-10-05  
PRIORITY APPLICATION NUMBER: 60/238,400  
PRIORITY FILING DATE: 2000-10-06  
PRIORITY APPLICATION NUMBER: 60/238,397  
PRIORITY FILING DATE: 2000-10-06  
PRIORITY APPLICATION NUMBER: 60/238,401  
PRIORITY FILING DATE: 2000-10-06  
PRIORITY APPLICATION NUMBER: 60/238,379  
PRIORITY FILING DATE: 2000-10-06  
PRIORITY APPLICATION NUMBER: 60/238,402  
PRIORITY FILING DATE: 2000-10-06  
PRIORITY APPLICATION NUMBER: 30/238,384  
PRIORITY FILING DATE: 2000-10-06  
PRIORITY APPLICATION NUMBER: 60/238,373  
PRIORITY FILING DATE: 2000-10-06  
PRIORITY APPLICATION NUMBER: 60/238,372  
PRIORITY FILING DATE: 2000-10-06  
PRIORITY APPLICATION NUMBER: 60/238,383  
PRIORITY FILING DATE: 2000-10-06  
PRIORITY APPLICATION NUMBER: 60/238,382  
PRIORITY FILING DATE: 2000-10-06  
PRIORITY APPLICATION NUMBER: 60/275,892  
PRIORITY FILING DATE: 2001-03-14  
PRIORITY APPLICATION NUMBER: 60/296,860  
PRIORITY FILING DATE: 2001-06-08  
NUMBER OF SEQ ID NOS: 200  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 39  
LENGTH: 1734  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (112)..(1686)  
US-10-096-625-39

Query Match 99.8%; Score 1574.8; DB 16; Length 1734;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1576; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCACCTTGTCCTCCGACGACGCTGTCCACGTCGCCACGCTCCGCTGTCCAAACAGTCC 60  
DB 112 ATGGCCACCTTGTCCTCCGACGACGCTGTCCACGTCGCCACGCTCCGCTGTCCAAACAGTCC 171  
QY 61 CAGGCCAAGATGAGCGGCAATGTTCCGACGATGGGTTTTCAGCGGCCACCGATGAGGAG 120  
DB 172 CAGGCCAAGATGAGCGGCAATGTTCCGACGATGGGTTTTCAGCGGCCACCGATGAGGAG 231  
QY 121 CGGGTGGGCTTCGCGATGTCGACGACCTCGACTTTGACGACCGCCAGGCGCTTCGACATG 180  
DB 232 CGGGTGGGCTTCGCGATGTCGACGACCTCGACTTTGACGACCGCCAGGCGCTTCGACATG 291  
QY 181 GACATCCTGAAAGCCGAGGAGAGCCCTCGCGGAGACGAGGGCGCTGAAAGCGGCCGTCGAG 240

292 GACATCCTGAAAGCCGAGGAGAGCCCTCGCGGAGACGAGGGCGCTGAAAGCGGCCGTCGAG 351  
241 GGAGACATTCATATCAGCGAGGAGCGGAGCTCCTCTGCGCGCCCTCCGGCTCCAAAGGAC 300  
352 GGAGACATTCATATCAGCGAGGAGCGGAGCTCCTCTGCGCGCCCTCCGGCTCCAAAGGAC 411  
301 CAGGTGGGAGGTGGTGGCGAATTCGGGGGCGACGACAAAGCCCAAAATACAGCGCTGGGAG 360  
412 CAGGTGGGAGGTGGTGGCGAATTCGGGGGCGACGACAAAGCCCAAAATACAGCGCTGGGAG 471  
361 GCAGGCTGGACGTCGACCAAGCCCATCAGGGCATGTTGCTGCTGGGCGCTTACCCTACGCC 420  
472 GCAGGCTGGACGTCGACCAAGCCCATCAGGGCATGTTGCTGCTGGGCGCTTACCCTACGCC 531  
421 ATCTGTGACGCGGCTACCTGGGGTGTCTCATCATCTTTCGCGCGCGTGTGTGTGTCG 480  
532 ATCTGTGACGCGGCTACCTGGGGTGTCTCATCATCTTTCGCGCGCGTGTGTGTGTCG 591  
481 TACACCGGCAAGATCCTCATCGCTGCTGTACGAGGAGAAATGAAGCGGCGAGGTGTG 540  
592 TACACCGGCAAGATCCTCATCGCTGCTGTACGAGGAGAAATGAAGCGGCGAGGTGTG 651  
541 CGCGTGGGAGCTGTCAGTGGCCATAGCCAAAGCCCTGCTGCGCGCGCTTCCCAACG 600  
652 CGCGTGGGAGCTGTCAGTGGCCATAGCCAAAGCCCTGCTGCGCGCGCTTCCCAACG 711  
601 CTGGCGCGCGAGTGGTGAACGTAGCGCAGATCATCGAGCTGGTGATGACGTGATCCTG 660  
712 CTGGCGCGCGAGTGGTGAACGTAGCGCAGATCATCGAGCTGGTGATGACGTGATCCTG 771  
661 TACGTGTGTGTGAGTGGCAACCTCATGTATCAACAGCTTCCGCGGCTCCCGTGTGCGAG 720  
772 TACGTGTGTGTGAGTGGCAACCTCATGTATCAACAGCTTCCGCGGCTCCCGTGTGCGAG 831  
721 AAGTCTGTGTCCATATCGCCACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780  
832 AAGTCTGTGTCCATATCGCCACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 891  
781 AAGCGCGTGTCCAAAGTTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840  
892 AAGCGCGTGTCCAAAGTTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 951  
841 GTCATAGCTACTGTCTATCGCGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900  
952 GTCATAGCTACTGTCTATCGCGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1011  
901 ATCGACGTCAAGAGTTCCCATCTCCATTTGGCATCATCGTGTTCAGCTACAGCTCTCAG 960  
1012 ATCGACGTCAAGAGTTCCCATCTCCATTTGGCATCATCGTGTTCAGCTACAGCTCTCAG 1071  
961 ATCTTCTGCTTCCGCTGAGGGCAATATGAGGAGCCAGCGAGTTCACCTGTCATGATG 1020  
1072 ATCTTCTGCTTCCGCTGAGGGCAATATGAGGAGCCAGCGAGTTCACCTGTCATGATG 1131  
1021 AACTGGACGACATCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080  
1132 AACTGGACGACATCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1191  
1081 ACCTGGGCGGAGAGACCAAGAGGTTCATCAGGATAACCTGCGCGCTCCATTCGCGGC 1140  
1192 ACCTGGGCGGAGAGACCAAGAGGTTCATCAGGATAACCTGCGCGCTCCATTCGCGGC 1251  
1141 GTGGTCAACATCTTCTGCTGCGCAAGCGCTGTTGCTCTCTCTGCTGCTGCTGCTGCTG 1200  
1252 GTGGTCAACATCTTCTGCTGCGCAAGCGCTGTTGCTCTCTCTGCTGCTGCTGCTGCTG 1311  
1201 GCTGTGAGGCTGCTGAGAGTGTGCTTTCAGGAGGAGCGAGCGCGCTTTTTCGCGGC 1260  
1312 GCTGTGAGGCTGCTGAGAGTGTGCTTTCAGGAGGAGCGAGCGCGCTTTTTCGCGGC 1371  
1261 TGTTCAGGCGGAGCGGCGCTGAGTCTCTGGGGGCTGACGCTGCGCTGCGGCTGCGGCTGCTC 1320

;; PRIOR FILING DATE: 2000-10-06  
;; PRIOR APPLICATION NUMBER: 60/275,892  
;; PRIOR FILING DATE: 2001-03-14  
;; PRIOR APPLICATION NUMBER: 60/296,860  
;; PRIOR FILING DATE: 2001-06-08  
;; NUMBER OF SEQ ID NOS: 200  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 37  
;; LENGTH: 1734  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: (112)..(1686)  
US-10-096-625-37

Query Match 99.8%; Score 1574.8; DB 16; Length 1734;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1576; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 ATGGCCACCTTGCTCCGACGACGAGCTGTCCAAAGTGGCCACGCTCCGTTGTCACCAAGTCC 60  
DB 112 ATGGCCACCTTGCTCCGACGACGAGCTGTCCAAAGTGGCCACGCTCCGTTGTCACCAAGTCC 171  
  
QY 61 CAGGCCAAGATGAGCGGCATGTTCCGCCAGATGGGTTTTCAGGCGGCCACGGATGAGGAG 120  
DB 172 CAGGCCAAGATGAGCGGCATGTTCCGCCAGATGGGTTTTCAGGCGGCCACGGATGAGGAG 231  
  
QY 121 GCGGTGGGCTTCGCGCATTCGACGACCTTCGACTTTGAGCACCGCCAGGSCCTTCGACATG 180  
DB 232 GCGGTGGGCTTCGCGCATTCGACGACCTTCGACTTTGAGCACCGCCAGGSCCTTCGACATG 291  
  
QY 181 GACATCTGTAAGCCGAGGAGAGCCCTCGGGGACGAGGCGCTGAAGCGCCGCTCGAG 240  
DB 292 GACATCTGTAAGCCGAGGAGAGCCCTCGGGGACGAGGCGCTGAAGCGCCGCTCGAG 351  
  
QY 241 GGAGACATCATTTATCAGCGAGGAGCGGAGCTCTCTGCGCCCTCCGCTCCAAAGGAC 300  
DB 352 GGAGACATCATTTATCAGCGAGGAGCGGAGCTCTCTGCGCCCTCCGCTCCAAAGGAC 411  
  
QY 301 CAGGTGGAGGTGGTGGCGAATTCGGGGGCCACGACAAAGCCCAAAATTCAGCGGTGGGAG 360  
DB 412 CAGGTGGAGGTGGTGGCGAATTCGGGGGCCACGACAAAGCCCAAAATTCAGCGGTGGGAG 471  
  
QY 361 GCAGGTGGACCTGACCAACGACATCCAGGCGATGTTGCTGCTGGGCTTACCTTAGCC 420  
DB 472 GCAGGTGGACCTGACCAACGACATCCAGGCGATGTTGCTGCTGGGCTTACCTTAGCC 531  
  
QY 421 ATCTGTCACGCGGCTACCTGGGTTGTTTCTCATCATCTTCGCGCGCTGTTGCTGC 480  
DB 532 ATCTGTCACGCGGCTACCTGGGTTGTTTCTCATCATCTTCGCGCGCTGTTGCTGC 591  
  
QY 481 TACACGGCAAGATCTCTACGCGTGCCTGTACGAGGAGAAATGAAGCGCGAGGTGGTG 540  
DB 592 TACACGGCAAGATCTCTACGCGTGCCTGTACGAGGAGAAATGAAGCGCGAGGTGGTG 651  
  
QY 541 CCGGTGGGACCTGCTAGTGGGCAATAGCAAGCCCTGTGCGCCCGCGCTTCCCAAGC 600  
DB 652 CCGGTGGGACCTGCTAGTGGGCAATAGCAAGCCCTGTGCGCCCGCGCTTCCCAAGC 711  
  
QY 601 CTGGGCGGCGAGTGGTGAACGTAGCGAGATCATCGAGCTGGTGTGATGAGTGCATCC 660  
DB 712 CTGGGCGGCGAGTGGTGAACGTAGCGAGATCATCGAGCTGGTGTGATGAGTGCATCC 771  
  
QY 661 TACGTGGTGGTGGCAACCTCATGTACCAAGCTTCCCGGGGCTGCCCGTGTGCGCAG 720  
DB 772 TACGTGGTGGTGGCAACCTCATGTACCAAGCTTCCCGGGGCTGCCCGTGTGCGCAG 831  
  
QY 721 AAGTCTGTGTCATTATCGCACGGCGGTGCTGCTGCTGCGCTTCCCTTAAGAACCCTC 780  
DB 832 AAGTCTGTGTCATTATCGCACGGCGGTGCTGCTGCTGCGCTTCCCTTAAGAACCCTC 891  
  
QY 781 AAGGCCGTGTCCAAAGTTTCAGTCTGCTGTGCACTCTGGCCGACCTTCGTCATCAATATCCTG 840

DB 892 AAGCCGTGTCCAAAGTTTCAGTCTGCTGTGCACTCTGGCCCACTTCGTCATCAATATCCTG 951  
QY 841 GTCATAGCCTACTGTCTATCGCGGGCGCGCACTGGGCTTGGAGAAAGTCAAGTTCTAC 900  
DB 952 GTCATAGCCTACTGTCTATCGCGGGCGCGCACTGGGCTTGGAGAAAGTCAAGTTCTAC 1011  
QY 901 ATCGAGCTCAAGAAAGTTCCCATCTCCATTTGGGCATCATCGTGTTCAGCTACAGTCTCAG 960  
DB 1012 ATCGAGCTCAAGAAAGTTCCCATCTCCATTTGGGCATCATCGTGTTCAGCTACAGTCTCAG 1071  
QY 961 ATCTTCTGCTCTTCTGCTGAGGGAATATGACAGAGCCAGCGAGTTCCTCACTGCATGATG 1020  
DB 1072 ATCTTCTGCTCTTCTGCTGAGGGAATATGACAGAGCCAGCGAGTTCCTCACTGCATGATG 1131  
QY 1021 AACTGACGACATCGCAGCTCTGCTCAAGGGCTCTTTCGCGCTCGTGCCTACCTC 1080  
DB 1132 AACTGACGACATCGCAGCTCTGCTCAAGGGCTCTTTCGCGCTCGTGCCTACCTC 1191  
QY 1081 ACCTGGGCGACGAGACCAAGAGGTCATCACGGATAAACCCTGCGGCTCCATCCGGGCC 1140  
DB 1192 ACCTGGGCGACGAGACCAAGAGGTCATCACGGATAAACCCTGCGGCTCCATCCGGGCC 1251  
QY 1141 GTGCTCAAACATCTTCTGCTGGCCAAAGGCGCTGTTGCTTATCTCTGCGCATTTCTTGCC 1200  
DB 1252 GTGCTCAAACATCTTCTGCTGGCCAAAGGCGCTGTTGCTTATCTCTGCCATTTCTTGCC 1311  
QY 1201 GCTGTCCAGGTGCTGAGAAAGTCGCTCTTTCAGGAAAGGACGCGCGCTTTTCCGGGCC 1260  
DB 1312 GCTGTCCAGGTGCTGAGAAAGTCGCTCTTTCAGGAAAGGACGCGCGCTTTTCCGGGCC 1371  
QY 1261 TGCTACAGCGGACGCGGCGCTGAAAGTCCTGGGGGCTGACGCTGCGCTGCGGCTGCTC 1320  
DB 1372 TGCTACAGCGGACGCGGCGCTGAAAGTCCTGGGGGCTGACGCTGCGCTGCGGCTGCTC 1431  
QY 1321 GTCTTCCAGCTGCTCATGCGCATTTATGTGCGGCATCTCGCGCTGCTCATGGGCTCACC 1380  
DB 1432 GTCTTCCAGCTGCTCATGCGCATTTATGTGCGGCATCTCGCGCTGCTCATGGGCTCACC 1491  
QY 1381 GGAGCCTCAAGCGGCGCGGCTCTGTTTCTTGTGCGGAGCTCTTTCACCTGCGGCTG 1440  
DB 1492 GGAGCCTCAAGCGGCGCGGCTCTGTTTCTTGTGCGGAGCTCTTTCACCTGCGGCTG 1551  
QY 1441 CTCTGGCGCAAGCTGCTGCGGCAACCAAGTCTTTCGAGCTGCGCATCTCTGTCATCGGC 1500  
DB 1552 CTCTGGCGCAAGCTGCTGCGGCAACCAAGTCTTTCGAGCTGCGCATCTCTGTCATCGGC 1611  
QY 1501 GGAGCCTCAAGCGGCTGCGGCTTCTGTCGCTTCTGTCGAGGCTCATCGAAGCTTACCGA 1560  
DB 1612 GGAGCCTCAAGCGGCTGCGGCTTCTGTCGCTTCTGTCGAGGCTCATCGAAGCTTACCGA 1671  
QY 1561 ACCAAGCGGAGGACTAG 1578  
DB 1672 ACCAAGCGGAGGACTAG 1689

RESULT 15  
US-10-096-625-39  
; Sequence 39, Application US/10096625  
; Publication No. US20040068095A1  
; GENERAL INFORMATION:  
; APPLICANT: Shimkets, Richard A  
; APPLICANT: Taupier Jr, Raymond J  
; APPLICANT: Burgess, Catherine E  
; APPLICANT: Zerhusen, Bryan D  
; APPLICANT: Mezes, Peter S  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Malyankar, Uriel M  
; APPLICANT: Grosse, William M  
; APPLICANT: Alsbrook II, John P  
; APPLICANT: Lepley, Denise M  
; APPLICANT: Spytek, Kimberly Ann  
; APPLICANT: Li, Li



Qy	301	CAGGTGGAGGTGGTGGCGAAATTGGGGGCGACGACAAGCCCAAAATACAGCGCTGGGAG	360
Db	412	CAGGTGGAGGTGGTGGCGAAATTGGGGGCGACGACAAGCCCAAAATACAGCGCTGGGAG	471
Qy	361	GCAGGCTGGAAAGTGAACGCAACGCCATCAGAGGCATGTCGTGCTGGGCCCTACCCCTACGCC	420
Db	472	GCAGGCTGGAAAGTGAACGCAACGCCATCAGAGGCATGTCGTGCTGGGCCCTACCCCTACGCC	531
Qy	421	ATCCTTGCA CGCGCGCTACCTTGGGGTGTGTTTCTCATCATCTTGGCGCGCTGTGTGCTGCTG	480
Db	532	ATCCTTGCA CGCGCGCTACCTTGGGGTGTGTTTCTCATCATCTTGGCGCGCTGTGTGCTGCTG	591
Qy	481	TACACCGGCAAGATCTCATCTGCGGTGCTGTACGAGGAGAAATGAAGACGCGCAGGTGGTG	540
Db	592	TACACCGGCAAGATCTCATCTGCGGTGCTGTACGAGGAGAAATGAAGACGCGCAGGTGGTG	651
Qy	541	CGCGTGGGACCTGCTAGCTGGCCATAGCCAAAGCCCTGTCGGCCCCGGCTTCCCAACG	600
Db	652	CGCGTGGGACCTGCTAGCGGGCCATAGCCAAAGCCCTGTCGGCCCCGGCTTCCCAACG	711
Qy	601	CTGGCGGCGCGAGTGGTGAACGTAGGCGCAGATCATCGAGCTGGTGTGATGACGTGCATCTTG	660
Db	712	CTGGCGGCGCGAGTGGTGAACGTAGGCGCAGATCATCGAGCTGGTGTGATGACGTGCATCTTG	771
Qy	661	TACGTGGTGTGAGTGGCAACCTCATGTACAACAGCTTCCGGGGGTGCCCGTGTGCGAG	720
Db	772	TACGTGGTGTGAGTGGCAACCTCATGTACAACAGCTTCCGGGGGTGCCCGTGTGCGAG	831
Qy	721	AAGTCTGGTTCATATTACGCAAGCGCGTGTCTGCTGCTGCGCTTCCCTTAAGAACCTTC	780
Db	832	AAGTCTGGTTCATATTACGCAAGCGCGTGTCTGCTGCTGCGCTTCCCTTAAGAACCTTC	891
Qy	781	AAGCGCGTGTCCAAGTTTCAGTCTGCTGTGACATCTGGCCCCACATTCGTTCATCAATATCTTG	840
Db	892	AAGCGCGTGTCCAAGTTTCAGTCTGCTGTGACATCTGGCCCCACATTCGTTCATCAATATCTTG	951
Qy	841	GTCAATAGCCTACTGTCTATCGCGGGGCGCGAATGGGCGCTGGAGAAAGTCAAGTTCTAC	900
Db	952	GTCAATAGCCTACTGTCTATCGCGGGCGCGAATGGGCGCTGGAGAAAGTCAAGTTCTAC	1011
Qy	901	ATCGAGCTCAAGAAAGTTCCCATCTCCATTGGGCATCATCGTGTTCAGCTACAGCTCTCAG	960
Db	1012	ATCGAGCTCAAGAAAGTTCCCATCTCCATTGGGCATCATCGTGTTCAGCTACAGCTCTCAG	1071
Qy	961	ATCTTCTGCTTTCGCTGGAGGGCAATATGACAGAGCCAGCGAGTTCCTACTGCATGATG	1020
Db	1072	ATCTTCTGCTTTCGCTGGAGGGCAATATGACAGAGCCAGCGAGTTCCTACTGCATGATG	1131
Qy	1021	AACTGGACGACATCGCAGCTCGTGTCTCAAGGGCCCTTTCGCGCTCGTGCCTACCTC	1080
Db	1132	AACTGGACGACATCGCAGCTCGTGTCTCAAGGGCCCTTTCGCGCTCGTGCCTACCTC	1191
Qy	1081	ACCTGGGCGGACGAGACGAGGAGGTCAACGAGATAACCTGCCGGCTCCATCCGGGCC	1140
Db	1192	ACCTGGGCGGACGAGACGAGGAGGCCATACGAGATAACCTGCCGGCTCCATCCGGGCC	1251
Qy	1141	GTGCTCAACATCTTCTGTGTGGCCAAAGCGCTGTGCTTCTATCTCTGCGCATCTTTGTC	1200
Db	1252	GTGCTCAACATCTTCTGTGTGGCCAAAGCGCTGTGCTTCTATCTCTGCGCATCTTTGTC	1311
Qy	1201	GCTGTCGAGGTGCTGGAGAGTTCGCTTCCAGGAAGCGAGCGCGCTTTTTCCCGGCC	1260
Db	1312	GCTGTCGAGGTGCTGGAGAGTTCGCTTCCAGGAAGCGAGCGCGCTTTTTCCCGGCC	1371
Qy	1261	TGCTACAGCGCGACGCGCGCTGAAATCTTGGGGGTGACGCTGCGCTGGCGCTCGTC	1320
Db	1372	TGCTACAGCGCGACGCGCGCTTGAAGTCTCTGGGGGTGACGCTGCGCTGGCGCTCGTC	1431
Qy	1321	GTCTTTCAGCTGCTCATGGCCATTATGTGCGGCACATTCGCGCTGCTCATGGGCTCAC	1380
Db	1432	GTCTTTCAGCTGCTCATGGCCATTATGTGCGGCACATTCGCGCTGCTCATGGGCTCAC	1491
Qy	1381	GGCAGCCTCACGGGCGCGCGCTCTGTGTTTCTTGTGCTGCGCAGCCCTTTTCACTGCGCCTG	1440

Qy 901 ATCGACGTCAGAAAGTTCCCATCTCCCATCTCCATTTGGCATCATCGTGTTCAGTACACGCTCTCAG 960  
Db 1012 ATCGACGTCAGAAAGTTCCCATCTCCCATCTCCATTTGGCATCATCGTGTTCAGTACACGCTCTCAG 1071  
Qy 961 ATCTTCTGCTTCCGCTGGAGGGCAATATGACGAGCCGACGAGTTCCACTGCAATGATG 1020  
Db 1072 ATCTTCTGCTTCCGCTGGAGGGCAATATGACGAGCCGACGAGTTCCACTGCAATGATG 1131  
Qy 1021 AACTGGACGCATCGCAGCGCTGGTCTCAAGGGCTCTTCGGGCTCGTGGCTACTCTC 1080  
Db 1132 AACTGGACGCATCGCAGCGCTGGTCTCAAGGGCTCTTCGGGCTCGTGGCTACTCTC 1191  
Qy 1081 ACCTGGGCGCAGCAGACCAAGAGGCTCATACGATAACTTCCCGGCTCCATCCGGGCC 1140  
Db 1192 ACCTGGGCGCAGCAGACCAAGAGGCTCATACGATAACTTCCCGGCTCCATCCGGGCC 1251  
Qy 1141 GTGCTCAACATCTTTCTGGTGGCCAAAGCGCTGTGTCTTATCTCTGCAATCTTTTGGCC 1200  
Db 1252 GTGCTCAACATCTTTCTGGTGGCCAAAGCGCTGTGTCTTATCTCTGCAATCTTTTGGCC 1311  
Qy 1201 GCTGTCAGGTGCTGGAGAGTGCCTCTTCCAGGAAGCAGCGCGCTTTTCCCGGCC 1260  
Db 1312 GCTGTCAGGTGCTGGAGAGTGCCTCTTCCAGGAAGCAGCGCGCTTTTCCCGGCC 1371  
Qy 1261 TGCTACAGCGCGACGGCGCTGAAGTCTTGGGGCTGACGCTGCGCTGCGCTCGTC 1320  
Db 1372 TGCTACAGCGCGACGGCGCTGAAGTCTTGGGGCTGACGCTGCGCTGCGCTCGTC 1431  
Qy 1321 GTCTTACGCTGCTCATGGCCATTTATGTGCCGACCTTTCGCGCTGCTCATGGGCTCAAC 1380  
Db 1432 GTCTTACGCTGCTCATGGCCATTTATGTGCCGACCTTTCGCGCTGCTCATGGGCTCAAC 1491  
Qy 1381 GGCAGCCTCAGGGCGCGGCTCTGTTCTTGTGTCGCGACCTCTTTCACCTGCGGCTG 1440  
Db 1492 GGCAGCCTCAGGGCGCGGCTCTGTTCTTGTGTCGCGACCTCTTTCACCTGCGGCTG 1551  
Qy 1441 CTCTGGCGCAAGCTGCTGTGGCAACCAAGTCTTCTTGCAGCTCGCAGCTCTTTCGTCATCGGC 1500  
Db 1552 CTCTGGCGCAAGCTGCTGTGGCAACCAAGTCTTCTTGCAGCTCGCAGCTCTTTCGTCATCGGC 1611  
Qy 1501 GGCATCTGACGCGTCTCGGCTTGTGTCACCTCCCTCGAGGGCTCATCGAAGGCTACCGA 1560  
Db 1612 GGCATCTGACGCGTCTCGGCTTGTGTCACCTCCCTCGAGGGCTCATCGAAGGCTACCGA 1671  
Qy 1561 ACCAACCGGAGGACTAG 1578  
Db 1672 ACCAACCGGAGGACTAG 1689

RESULT 13

US-10-096-625-35  
; Sequence 35, Application US/10096625  
; Publication No. US20040068095A1  
; GENERAL INFORMATION:  
; APPLICANT: Shimkets, Richard A  
; APPLICANT: Taupier Jr, Raymond J  
; APPLICANT: Burgess, Catherine E  
; APPLICANT: Zerhusen, Bryan D  
; APPLICANT: Mezes, Peter S  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Malyankar, Uriel M  
; APPLICANT: Grosse, William M  
; APPLICANT: Alsobrook II, John P  
; APPLICANT: Lepley, Denise M  
; APPLICANT: Spytek, Kimberly Ann  
; APPLICANT: Li, Li  
; APPLICANT: Edinger, Shlomit  
; APPLICANT: Gerlach, Valerie  
; APPLICANT: Ellerman, Karen  
; APPLICANT: MacDougall, John R  
; APPLICANT: Gunther, Erik  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Stone, David J

; APPLICANT: Smithson, Glennda  
; APPLICANT: Szekeres Jr, Edward S  
; APPLICANT: Ji, Weizhen  
; TITLE OF INVENTION: No. US20040068095A1el Human Proteins, Polynucleotides Encoding T  
; TITLE OF INVENTION: Methods Of Using The Same  
; FILE REFERENCE: 21402-141 CIP  
; CURRENT APPLICATION NUMBER: US/10/096,625  
; CURRENT FILING DATE: 2002-03-13  
; PRIOR APPLICATION NUMBER: 09/972,211  
; PRIOR FILING DATE: 2001-10-05  
; PRIOR APPLICATION NUMBER: 60/238,325  
; PRIOR FILING DATE: 2000-10-05  
; PRIOR APPLICATION NUMBER: 60/238,323  
; PRIOR FILING DATE: 2000-10-05  
; PRIOR APPLICATION NUMBER: 60/238,400  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/238,397  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/238,401  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/238,379  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/238,402  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 30/238,384  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/238,373  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/238,372  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/238,383  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/238,382  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/275,892  
; PRIOR FILING DATE: 2001-03-14  
; PRIOR APPLICATION NUMBER: 60/296,860  
; PRIOR FILING DATE: 2001-06-08  
; NUMBER OF SEQ ID NOS: 200  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 35  
; LENGTH: 1734  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (112)..(1686)  
US-10-096-625-35

Query Match 99.8%; Score 1574.8; DB 16; Length 1734;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1576; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGGCCACCTTGCTCGCAGCAAGCTGTCCAACTGGCGCACGTCCTGTGTCCAAACAAGTCC 60  
Db 112 ATGGCCACCTTGCTCGCAGCAAGCTGTCCAACTGGCGCACGTCCTGTGTCCAAACAAGTCC 171  
Qy 61 CAGGCCAAGATGAGCGGCATGTTTCGCCAGGATGGGTTTTTCAGCGCGCCACGGATGAGGAG 120  
Db 172 CAGGCCAAGATGAGCGGCATGTTTCGCCAGGATGGGTTTTTCAGCGCGCCACGGATGAGGAG 231  
Qy 121 GCGGTGGGCTTTCGCGCATTTGCCAGACCTTCGATTTGAGCACCGCCAGGGCTTCGAGATG 180  
Db 232 GCGGTGGGCTTTCGCGCATTTGCCAGACCTTCGATTTGAGCACCGCCAGGGCTTCGAGATG 291  
Qy 181 GACATCTGAAAGCCGAGGAGAGCCCTCGCGGAGAGAGCGCTCAAGCGCCCGCTCGAG 240  
Db 292 GACATCTGAAAGCCGAGGAGAGCCCTCGCGGAGAGAGCGCTCAAGCGCCCGCTCGAG 351  
Qy 241 GGAGACATCATTTATCAGCGAGGACGCGAGTCTCTCTCCGCGCTTCGCGCTCCAAAGGAC 300  
Db 352 GGAGACATCATTTATCAGCGAGGACGCGAGTCTCTCTCCGCGCTTCGCGCTCCAAAGGAC 411

QY 1501 GGCACTCTGAGGCTGCGGCTTCGGTGCACCTCCCTGAGGGGCTCATCGAAGCTTACCGA 1560  
Db |||||||  
QY 1561 ACCAACGGGAGGACTAG 1578  
Db |||||||  
1672 ACCAACGGGAGGACTAG 1689

RESULT 12  
US-10-096-625-31  
; Sequence 31, Application US/10096625  
; Publication No. US20040068095A1  
; GENERAL INFORMATION:  
; APPLICANT: Shimkets, Richard A  
; APPLICANT: Taupier Jr, Raymond J  
; APPLICANT: Burgess, Catherine E  
; APPLICANT: Zerhusen, Bryan D  
; APPLICANT: Mezes, Peter S  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Malyankar, Uriel M  
; APPLICANT: Grosse, William M  
; APPLICANT: Alsobrook II, John P  
; APPLICANT: Lepley, Denise M  
; APPLICANT: Spytek, Kimberly Ann  
; APPLICANT: Li, Li  
; APPLICANT: Edinger, Shlomit  
; APPLICANT: Gerlach, Valerie  
; APPLICANT: Ellerman, Karen  
; APPLICANT: MacDougall, John R  
; APPLICANT: Gunther, Erik  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Stone, David J  
; APPLICANT: Smithson, Glenda  
; APPLICANT: Szekeres Jr, Edward S  
; APPLICANT: Ji, Weizhen  
; TITLE OF INVENTION: No. US20040068095A1el Human Proteins, Polynucleotides Encoding Th  
; TITLE OF INVENTION: Methods of Using The Same  
; FILE REFERENCE: 21402-141 CIP  
; CURRENT APPLICATION NUMBER: US/10/096,625  
; CURRENT FILING DATE: 2002-03-13  
; PRIOR APPLICATION NUMBER: 09/972,211  
; PRIOR FILING DATE: 2001-10-05  
; PRIOR APPLICATION NUMBER: 60/238,325  
; PRIOR FILING DATE: 2000-10-05  
; PRIOR APPLICATION NUMBER: 60/238,323  
; PRIOR FILING DATE: 2000-10-05  
; PRIOR APPLICATION NUMBER: 60/238,400  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/238,397  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/238,401  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/238,379  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/238,402  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 30/238,384  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/238,373  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/238,372  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/238,383  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/238,382  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/275,892  
; PRIOR FILING DATE: 2001-03-14  
; PRIOR APPLICATION NUMBER: 60/296,860  
; PRIOR FILING DATE: 2001-06-08  
; NUMBER OF SEQ ID NOS: 200  
; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 31  
; LENGTH: 1734  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (112)..(1686)  
US-10-096-625-31  
  
Query Match 99.8%; Score 1574.8; DB 16; Length 1734;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1576; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 ATGGCCACCTTGTCTCCGCGAGCAAGCTGTCCAAACGTGGCCACGCTCCGTGTCCAAACAAAGTCC 60  
Db |||||||  
QY 112 ATGGCCACCTTGTCTCCGCGAGCAAGCTGTCCAAACGTGGCCACGCTCCGTGTCCAAACAAAGTCC 171  
Db |||||||  
QY 61 CAGGCCAAGATGAGCGGCATGTTGCCAGGATGGGTTTTTCAGGGGGCCACGGATGAGGAG 120  
Db |||||||  
QY 172 CAGGCCAAGATGAGCGGCATGTTGCCAGGATGGGTTTTTCAGGGGGCCACGGATGAGGAG 231  
Db |||||||  
QY 121 CGGTGGGCTTTCGGCATTGCGACACCTTCGACATTTGAGCACCCGCGGCTTCAGATG 180  
Db |||||||  
QY 232 CGGTGGGCTTTCGGCATTGCGACACCTTCGACATTTGAGCACCCGCGGCTTCAGATG 291  
Db |||||||  
QY 181 GACATCTGAAACCGGAGGAGAGCCCTGCGGGGACGAGGGCGCTGAAAGCGCCCGTCCGAG 240  
Db |||||||  
QY 292 GACATCTGAAACCGGAGGAGAGCCCTGCGGGGACGAGGGCGCTGAAAGCGCCCGTCCGAG 351  
Db |||||||  
QY 241 GGAGACATCCATTATACGAGGAGCGGAGCTCTCTGCGGCGCTCCGCGTCCAAAGGAC 300  
Db |||||||  
QY 352 GGAGACATCCATTATACGAGGAGCGGAGCTCTCTGCGGCGCTCCGCGTCCAAAGGAC 411  
Db |||||||  
QY 301 CAGGTGGGAGGTGGTGGGGAATTCGGGGGCGACGACAGCCCAAAATCACGGCGTGGGAG 360  
Db |||||||  
QY 412 CAGGTGGGAGGTGGTGGGGAATTCGGGGGCGACGACAGCCCAAAATCACGGCGTGGGAG 471  
Db |||||||  
QY 361 CGAGGCTCGAAGCTGACCAACGCCATCCAGGGCATGTTCTGCTGGGCGCTACCTACGCC 420  
Db |||||||  
QY 472 CGAGGCTCGAAGCTGACCAACGCCATCCAGGGCATGTTCTGCTGGGCGCTACCTACGCC 531  
Db |||||||  
QY 421 ATCTTGCAACGGCGGTACTCTGGGGTGTGTTCTCATCATCTTTCGCGCGCTGTGTGCTGC 480  
Db |||||||  
QY 532 ATCTTGCAACGGCGGTACTCTGGGGTGTGTTCTCATCATCTTTCGCGCGCTGTGTGCTGC 591  
Db |||||||  
QY 481 TACACGGGCAAGATCCTCATCGCGTGTCTGACGAGGAGATGAAAGCGGCGAGGTGGT 540  
Db |||||||  
QY 592 TACACGGGCAAGATCCTCATCGCGTGTCTGACGAGGAGATGAAAGCGGCGAGGTGGT 651  
Db |||||||  
QY 541 CGCGTGGCGGACTCGTACGTGGCCATAGCCAAACGCTGCTGCGCGCGCGCTTCCCAACG 600  
Db |||||||  
QY 652 CGCGTGGCGGACTCGTACGTGGCCATAGCCAAACGCTGCTGCGCGCGCGCTTCCCAACG 711  
Db |||||||  
QY 601 TTGGCGCGCGAGTGGTGAACGTAGCGAGATCATCGAGCTGGTGTATGACGTGCTCTG 660  
Db |||||||  
QY 712 TTGGCGCGCGAGTGGTGAACGTAGCGAGATCATCGAGCTGGTGTATGACGTGCTCTG 771  
Db |||||||  
QY 661 TACGTGGTGGTGGGCAACCTCATGTACACAGCTTCCCGGGGCTGCCGCTGCGCAG 720  
Db |||||||  
QY 772 TACGTGGTGGTGGGCAACCTCATGTACACAGCTTCCCGGGGCTGCCGCTGCGCAG 831  
Db |||||||  
QY 721 AAGTCTGTGTCATTATCGCCACGGCGGTGCTGCTGCTTTCCTTAAGAACCTC 780  
Db |||||||  
QY 832 AAGTCTGTGTCATTATCGCCACGGCGGTGCTGCTGCTTTCCTTAAGAACCTC 891  
Db |||||||  
QY 781 AAGGCCGTGTCCAAGTTCAAGTGTGTGACCTCTGGGCCCACTTCGTATCAATATCTCTG 840  
Db |||||||  
QY 892 AAGGCCGTGTCCAAGTTCAAGTGTGTGACCTCTGGGCCCACTTCGTATCAATATCTCTG 951  
Db |||||||  
QY 841 GTCATAGCCTACTGTCTATCGCGGGGCGGAGCTGGGCTGGGAGAGGTCAAGTTCTAC 900  
Db |||||||  
QY 952 GTCATAGCCTACTGTCTATCGCGGGGCGGAGCTGGGCTGGGAGAGGTCAAGTTCTAC 1011  
Db |||||||

```
; FILE REFERENCE: 21402-141
; CURRENT APPLICATION NUMBER: US/09/972,211
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 60/238,325
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/238,323
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/238,400
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,397
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,401
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,379
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,402
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 30/238,384
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,373
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,372
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,383
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,382
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/275,892
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/296,860
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 198
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 39
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (112)..(1686)
; US-09-972-211-39

Query Match          99.8%; Score 1574.8; DB 11; Length 1734;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1576; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGCCACCTTGCTCCGCGAGCAAGCTGTCCAACTGGCCAGCTCCGTTGCCAACCAAGTCC 60
Db 112 ATGCCACCTTGCTCCGCGAGCAAGCTGTCCAACTGGCCAGCTCCGTTGCCAACCAAGTCC 171

Qy 61 CAGGCCAAGATGAGCGGATGTTCCGCAAGATGGGTTTTCAGGCGGCCACGGATGAGGAG 120
Db 172 CAGGCCAAGATGAGCGGATGTTCCGCAAGATGGGTTTTCAGGCGGCCACGGATGAGGAG 231

Qy 121 GCGGTGGCTTCGCGCATTTGCGAGCACTCGACTTTGAGCACTCCAGGCGCTCCAGATG 180
Db 232 GCGGTGGCTTCGCGCATTTGCGAGCACTCGACTTTGAGCACTCCAGGCGCTCCAGATG 291

Qy 181 GACATCTCTGAAAGCCGAGGAGAGCCCTTCGCGGACGAGGCGCTGAAGCGCCCGTCGAG 240
Db 292 GACATCTCTGAAAGCCGAGGAGAGCCCTTCGCGGACGAGGCGCTGAAGCGCCCGTCGAG 351

Qy 241 GGAGACATCCATTATCAGCGAGGACGGAGCTCTCTGCGGCCCTCGGGTCCAAGGAC 300
Db 352 GGAGACATCCATTATCAGCGAGGACGGAGCTCTCTGCGGCCCTCGGGTCCAAGGAC 411

Qy 301 CAGGTGGAGGTGGTGGCGAATTCGGGGGCCACGACAAAGCCAAATCAGGGCTGGGAG 360
Db 412 CAGGTGGAGGTGGTGGCGAATTCGGGGGCCACGACAAAGCCAAATCAGGGCTGGGAG 471

Qy 361 GCAGGCTGGAACTGACCAAGCCATCCAGGGCATGTTCTGCTGGGCGCTACCCCTAGGCC 420
Db 472 GCAGGCTGGAACTGACCAAGCCATCCAGGGCATGTTCTGCTGGGCGCTACCCCTAGGCC 531
```

```
Qy 421 ATCTGTGCAACGCGCGCTACCTGGGGTTCGTTCTCATCATCTTCGCGCGGCTTGTGTGTCG 480
Db 532 ATCTGTGCAACGCGCGCTACCTGGGGTTCGTTCTCATCATCTTCGCGCGGCTTGTGTGTCG 591

Qy 481 TACACCGGCAAGATCCCTCATCGCGTGCCTGTACGAGGAGAAATGAAGACGGCGAGTGGTG 540
Db 592 TACACCGGCAAGATCCCTCATCGCGTGCCTGTACGAGGAGAAATGAAGACGGCGAGTGGTG 651

Qy 541 CGGTGCGGGAATCTGTAAGTGGCCATAGCAACGCTGTGCGCCCGCGCTTCCCAACG 600
Db 652 CGGTGCGGGAATCTGTAAGTGGCCATAGCAACGCTGTGCGCCCGCGCTTCCCAACG 711

Qy 601 CTGGGCGCGGAGTGGTGAAGTAGCCAGATCATCGAGCTGTGATGACGTGATCCTG 660
Db 712 CTGGGCGCGGAGTGGTGAAGTAGCCAGATCATCGAGCTGTGATGACGTGATCCTG 771

Qy 661 TACGTGTGTGTGAGTGGCAACCTCATGTACAAAGCTTCCCGGGGCTGCCGTGTGCGAG 720
Db 772 TACGTGTGTGTGAGTGGCAACCTCATGTACAAAGCTTCCCGGGGCTGCCGTGTGCGAG 831

Qy 721 AAGTCTGTGTCCATTATTCGCAACGCGCTGTGCTGCTTTCGCGCTTCTTAAAGAACTC 780
Db 832 AAGTCTGTGTCCATTATTCGCAACGCGCTGTGCTGCTTTCGCGCTTCTTAAAGAACTC 891

Qy 781 AAGGCGGTGTCGAAGTTCAGTCTGCTGTGCACTCTGCGCCCACTTCGTCAATATCCTG 840
Db 892 AAGGCGGTGTCGAAGTTCAGTCTGCTGTGCACTCTGCGCCCACTTCGTCAATATCCTG 951

Qy 841 GTCATAGCTTACTGTCTATCGCGGCGCGAGCTGGGCTTGGGAGAAAGTCAAGTTCCTAC 900
Db 952 GTCATAGCTTACTGTCTATCGCGGCGCGAGCTGGGCTTGGGAGAAAGTCAAGTTCCTAC 1011

Qy 901 ATCGAGCTCAAGAAAGTTCCTCATTTGGCATCATCGTGTTCAGTCAACGCTCTCAG 960
Db 1012 ATCGAGCTCAAGAAAGTTCCTCATTTGGCATCATCGTGTTCAGTCAACGCTCTCAG 1071

Qy 961 ATCTTCTGCTGCTGCTGGAGGGCAATATGCAAGAGCCCAAGGAGTTCACATGATGATG 1020
Db 1072 ATCTTCTGCTGCTGCTGGAGGGCAATATGCAAGAGCCCAAGGAGTTCACATGATG 1131

Qy 1021 AACTGGACGACATCGCAGCTGCTCAAGGGCTCTTCGCGCTCGCTCGCTACCTC 1080
Db 1132 AACTGGACGACATCGCAGCTGCTCAAGGGCTCTTCGCGCTCGCTCGCTACCTC 1191

Qy 1081 ACTTGGGCGCAAGAGGAGTCAACGAGTAACTGCGCGGCTCCATCCCGGCTC 1140
Db 1192 ACTTGGGCGCAAGAGGAGTCAACGAGTAACTGCGCGGCTCCATCCCGGCTC 1251

Qy 1141 GTGGTCAACATCTTCTGGTGGCCAAAGGCGCTGTGCTCTATCTCTGCGCATTTCTTGGC 1200
Db 1252 GTGGTCAACATCTTCTGGTGGCCAAAGGCGCTGTGCTCTATCTCTGCGCATTTCTTGGC 1311

Qy 1201 GCTGTGAGGTGCTGAGAAAGTTCGCTTCCAGAAAGGAGCGCGGCTTTTCCCGGCTC 1260
Db 1312 GCTGTGAGGTGCTGAGAAAGTTCGCTTCCAGAAAGGAGCGCGGCTTTTCCCGGCTC 1371

Qy 1261 TGCTACAGCGGCAAGGCGCTGAAAGTCTCTGGGGGCTGACGCTGCGCTCGCGCTCGTC 1320
Db 1372 TGCTACAGCGGCAAGGCGCTGAAAGTCTCTGGGGGCTGACGCTGCGCTCGCGCTCGTC 1431

Qy 1321 GTCTTCAAGCTGCTCATGGCCATTTATGTGCGGCACTTCGCGCTGCTCATGGGCTCAAC 1380
Db 1432 GTCTTCAAGCTGCTCATGGCCATTTATGTGCGGCACTTCGCGCTGCTCATGGGCTCAAC 1491

Qy 1381 GGCAGCTCAAGGCGCGCGCTCTGTTTCTGCTGCGCAAGCTCTTCACTGCGGCTG 1440
Db 1492 GGCAGCTCAAGGCGCGCGCTCTGTTTCTGCTGCGCAAGCTCTTCACTGCGGCTG 1551

Qy 1441 CTCTGCGCAAGCTGCTGCGCAAGCTCTTCTGCAAGTCTTCTGCAAGTCTTCTGCGCTCG 1500
Db 1552 CTCTGCGCAAGCTGCTGCGCAAGCTCTTCTGCAAGTCTTCTGCGCTCGCTCGCTCGCTCG 1611
```

FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (112)...(1686)  
US-09-972-211-37

Query Match 99.8%; Score 1574.8; DB 11; Length 1734;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1576; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	1	ATGGCCACCTTCTCGGAGCAAGCTGTCACAGTGGCCACGTCGCTGTCACAGTCC	60
DB	112	ATGGCCACCTTCTCGGAGCAAGCTGTCACAGTGGCCACGTCGCTGTCACAGTCC	171
QY	61	CAGGCCAAGATCAGCGGATGTTCCGAGGATGGTTTTTCAGCGCGCCACAGGATCAGGAG	120
DB	172	CAGGCCAAGATCAGCGGATGTTCCGAGGATGGTTTTTCAGCGCGCCACAGGATCAGGAG	231
QY	121	GGGTGGGCTTCGCGCATTTGCGACGACCTCGACTTTGAGCACCGCCAGGGCCCTGCAGATG	180
DB	232	GGGTGGGCTTCGCGCATTTGCGACGACCTCGACTTTGAGCACCGCCAGGGCCCTGCAGATG	291
QY	181	GACATCTGAAGCCGAGGGAGAGCCCTCGGGGAGAGGGCGCTGAAGCGCCGCTCGAG	240
DB	292	GACATCTGAAGCCGAGGGAGAGCCCTCGGGGAGAGGGCGCTGAAGCGCCGCTCGAG	351
QY	241	GGAGACATCATATCAGCGAGCGAGCGAGCTCTCTGCGCGCCCTCGGGCTTCAAGGAC	300
DB	352	GGAGACATCATATCAGCGAGCGAGCGAGCTCTCTGCGCGCCCTCGGGCTTCAAGGAC	411
QY	301	CAGGTGGAGGTGGTGGCGAAATTCGGGGGCCACGACCAAGCCCAAAATCACGGCGTGGAG	360
DB	412	CAGGTGGAGGTGGTGGCGAAATTCGGGGGCCACGACCAAGCCCAAAATCACGGCGTGGAG	471
QY	361	CGAGGTGGAACGTGACCAACCGCATTCAGGGCATGTTGCTGCTGGGCTTACCTACGCC	420
DB	472	CGAGGTGGAACGTGACCAACCGCATTCAGGGCATGTTGCTGCTGGGCTTACCTACGCC	531
QY	421	ATCTCTCAGCGGGCTTACCTGGGTGTTTCTCATCATCTTCCGCGCGTGTGTGCTGC	480
DB	532	ATCTCTCAGCGGGCTTACCTGGGTGTTTCTCATCATCTTCCGCGCGTGTGTGCTGC	591
QY	481	TACACGGCAAGATCCTCATCGGTCCTGTATCAGGAGAAATGAAGCGCGAGGTGGTG	540
DB	592	TACACGGCAAGATCCTCATCGGTCCTGTATCAGGAGAAATGAAGCGCGAGGTGGTG	651
QY	541	CGGTCGGGACTCGTATCGTGGCATAGCCACGCTGTGCGCCCGCGCTTCCACG	600
DB	652	CGGTCGGGACTCGTATCGTGGCATAGCCACGCTGTGCGCCCGCGCTTCCACG	711
QY	601	CTGGGCGCGGAGTGGTGAACGTAGCGAGATCATCGAGCTGGTATGACGTGCATCTCG	660
DB	712	CTGGGCGCGGAGTGGTGAACGTAGCGAGATCATCGAGCTGGTATGACGTGCATCTCG	771
QY	661	TACGTGGTGGTGAAGTGGCAACCTCATGTACACAGCTTCCCGGGGTGCGCGTGTGCGAG	720
DB	772	TACGTGGTGGTGAAGTGGCAACCTCATGTACACAGCTTCCCGGGGTGCGCGTGTGCGAG	831
QY	721	AAGTCTGTGTCATATCGCCAGCGCGCTGCTGCTGCTGCTGCTGCTTAAAGAACCTTC	780
DB	832	AAGTCTGTGTCATATCGCCAGCGCGCTGCTGCTGCTGCTGCTTAAAGAACCTTC	891
QY	781	AGGCGCGTTCGAAGTTCAGTCTGCTGTCATCTGCGCCCACTTCGTGATCATATCTCG	840
DB	892	AGGCGCGTTCGAAGTTCAGTCTGCTGTCATCTGCGCCCACTTCGTGATCATATCTCG	951
QY	841	GTCAATAGCTTACTGTTCTATCGGGGCGCGAGCTGGGCGCTGGGAGAGGTCAAGTTCTAC	900
DB	952	GTCAATAGCTTACTGTTCTATCGGGGCGCGAGCTGGGCGCTGGGAGAGGTCAAGTTCTAC	1011
QY	901	ATCGAGCTCAAGAAGTTCCCATCTCCATTTGGCATCATCGTGTTCAGCTACAGCTCTAG	960
DB	1012	ATCGAGCTCAAGAAGTTCCCATCTCCATTTGGCATCATCGTGTTCAGCTACAGCTCTAG	1071

QY	961	ATCTTCTGCTTCGCTCGGAGGCAATATGACGAGCCCGAGGATTCACATGATGATG	1020
DB	1072	ATCTTCTGCTTCGCTCGGAGGCAATATGACGAGCCCGAGGATTCACATGATGATG	1131
QY	1021	AACGTGAGCGACATCGCAGCCTGCTGCTCAAGGGCTCTTTCGGCTGCTCGCTACCTC	1080
DB	1132	AACGTGAGCGACATCGCAGCCTGCTGCTCAAGGGCTCTTTCGGCTGCTCGCTACCTC	1191
QY	1081	ACCTGGGCGGACGAGACCAAGAGGTATACAGGATAACCTGCGCGGTCCATCCGCGCC	1140
DB	1192	ACCTGGGCGGACGAGACCAAGAGGTATACAGGATAACCTGCGCGGTCCATCCGCGCC	1251
QY	1141	GTGGTCAACATCTTTTCTGGTGCCCAAGCGCTGTGTCTATCTCTGCTGCTTCTTCC	1200
DB	1252	GTGGTCAACATCTTTTCTGGTGCCCAAGCGCTGTGTCTATCTCTGCTGCTTCTTCC	1311
QY	1201	GCTGTGAGGTGCTGGAGAGTCTCTTCCAGGAAGCGCGCTTTTCCCGGCC	1260
DB	1312	GCTGTGAGGTGCTGGAGAGTCTCTTCCAGGAAGCGCGCTTTTCCCGGCC	1371
QY	1261	TGCTACAGCGCGGACGGCGCTCTGAAGTCTCTGGGGCTGACGCTGCGCTGCGCTCGTC	1320
DB	1372	TGCTACAGCGCGGACGGCGCTCTGAAGTCTCTGGGGCTGACGCTGCGCTGCGCTCGTC	1431
QY	1321	GTCTTTCAGCTGCTCATGGCCATTTATGTGCGGCACTTTCGGCGTCTCATGGGCTCACC	1380
DB	1432	GTCTTTCAGCTGCTCATGGCCATTTATGTGCGGCACTTTCGGCGTCTCATGGGCTCACC	1491
QY	1381	GGCAGCTTCACGGGGCGCGCTCTGTCTTCTGCTGCGGCTCTTTCACCTGCGCTG	1440
DB	1492	GGCAGCTTCACGGGGCGCGCTCTGTCTTCTGCTGCGGCTCTTTCACCTGCGCTG	1551
QY	1441	CTCTGGCGCAAGCTGCTGTGCGCACCAAGTCTTCTTTCGACGCTCGCATCTTCTGTCATCGGC	1500
DB	1552	CTCTGGCGCAAGCTGCTGTGCGCACCAAGTCTTCTTTCGACGCTCGCATCTTCTGTCATCGGC	1611
QY	1501	GGCATCTGACAGCTGTCGGCTTCTGTCACCTCCCTCGAGGGCTCATCGAAGCTACCGA	1560
DB	1612	GGCATCTGACAGCTGTCGGCTTCTGTCACCTCCCTCGAGGGCTCATCGAAGCTACCGA	1671
QY	1561	ACCAAGCGGAGGACTAG 1578	
DB	1672	ACCAAGCGGAGGACTAG 1689	

RESULT 11

US-09-972-211-39  
; Sequence 39, Application US/09972211  
; Publication No. US20040048245A1  
; GENERAL INFORMATION:  
; APPLICANT: Shimkets, Richard A  
; APPLICANT: Taupier Jr, Raymond J  
; APPLICANT: Burgess, Catherine E  
; APPLICANT: Zerhusen, Bryan D  
; APPLICANT: Mezes, Peter S  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Malyankar, Uriel M  
; APPLICANT: Grosbe, William M  
; APPLICANT: Alsobrook II, John P  
; APPLICANT: Lepley, Denise M  
; APPLICANT: Spytek, Kimberly Ann  
; APPLICANT: Li, Li  
; APPLICANT: Edinger, Shlomit  
; APPLICANT: Gerlach, Valerie  
; APPLICANT: Ellerman, Karen  
; APPLICANT: MacDougall, John R  
; APPLICANT: Gunther, Erik  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Stone, David J  
; APPLICANT: Smithson, Glenda  
; APPLICANT: Szekeres Jr, Edward S  
; TITLE OF INVENTION: No. US20040048245A1el Human Proteins, Polynucleotides Encoding Ti  
; TITLE OF INVENTION: Methods Of Using The Same







Db 1600 GGCACTCTGCACGCTGTCGGCTTCGTGCACTCCCTCGAGGCGCTCATCGAAGCCCTACCGA 1659

Qy 1561 ACCAACCGGAGGACTAG 1578

Db 1660 ACCAACCGGAGGACTAG 1677

## RESULT 8

US-09-972-211-31

; Sequence 31, Application US/09972211

; Publication No. US20040048245A1

; GENERAL INFORMATION:

; APPLICANT: Shimkets, Richard A

; APPLICANT: Taupier Jr, Raymond J

; APPLICANT: Burgess, Catherine E

; APPLICANT: Zernusen, Bryan D

; APPLICANT: Mezes, Peter S

; APPLICANT: Rastelli, Luca

; APPLICANT: Malyankar, Uriel M

; APPLICANT: Grosse, William M

; APPLICANT: Alsobrook II, John P

; APPLICANT: Lepley, Denise M

; APPLICANT: Spytek, Kimberly Ann

; APPLICANT: Li, Li

; APPLICANT: Edinger, Shlomit

; APPLICANT: Gerlach, Valerie

; APPLICANT: Ellerman, Karen

; APPLICANT: Macdougall, John R

; APPLICANT: Gunther, Erik

; APPLICANT: Millet, Isabelle

; APPLICANT: Stone, David J

; APPLICANT: Smithson, Glennda

; APPLICANT: Szekeres Jr, Edward S

; TITLE OF INVENTION: No. US20040048245A1el Human Proteins, Polynucleotides Encoding Th

; FILE REFERENCE: 21402-141

; CURRENT APPLICATION NUMBER: US/09/972,211

; PRIOR FILING DATE: 2001-10-05

; PRIOR APPLICATION NUMBER: 60/238,325

; PRIOR FILING DATE: 2000-10-05

; PRIOR APPLICATION NUMBER: 60/238,323

; PRIOR FILING DATE: 2000-10-05

; PRIOR APPLICATION NUMBER: 60/238,400

; PRIOR FILING DATE: 2000-10-06

; PRIOR APPLICATION NUMBER: 60/238,397

; PRIOR FILING DATE: 2000-10-06

; PRIOR APPLICATION NUMBER: 60/238,401

; PRIOR FILING DATE: 2000-10-06

; PRIOR APPLICATION NUMBER: 60/238,379

; PRIOR FILING DATE: 2000-10-06

; PRIOR APPLICATION NUMBER: 60/238,402

; PRIOR FILING DATE: 2000-10-06

; PRIOR APPLICATION NUMBER: 60/238,384

; PRIOR FILING DATE: 2000-10-06

; PRIOR APPLICATION NUMBER: 60/238,373

; PRIOR FILING DATE: 2000-10-06

; PRIOR APPLICATION NUMBER: 60/238,372

; PRIOR FILING DATE: 2000-10-06

; PRIOR APPLICATION NUMBER: 60/238,383

; PRIOR FILING DATE: 2000-10-06

; PRIOR APPLICATION NUMBER: 60/238,382

; PRIOR FILING DATE: 2000-10-06

; PRIOR APPLICATION NUMBER: 60/275,892

; PRIOR FILING DATE: 2001-03-14

; PRIOR APPLICATION NUMBER: 60/296,860

; PRIOR FILING DATE: 2001-06-08

; NUMBER OF SEQ ID NOS: 198

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 31

; LENGTH: 1734

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (112)..(1686)

US-09-972-211-31

Query Match 99.8%; Score 1574.8; DB 11; Length 1734;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1576; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGGCCACCTTGCTCCGACGAAAGCTGTCCAAAGTGGCCACGTCGCTGTTCCAAAGTCC 60

Db 112 ATGGCCACCTTGCTCCGACGAAAGCTGTCCAAAGTGGCCACGTCGCTGTTCCAAAGTCC 171

Qy 61 CAGGCCAAGATGAGCGCATGTTCCGACGATGGTTCAGGGCCGACGATGAGGAG 120

Db 172 CAGGCCAAGATGAGCGCATGTTCCGACGATGGTTCAGGGCCGACGATGAGGAG 231

Qy 121 GCGGTGGGCTTCGCGCATTCGACGACCTCGACTTTTGGAGCAGCCAGGCGCTTCAGATG 180

Db 232 GCGGTGGGCTTCGCGCATTCGACGACCTCGACTTTTGGAGCAGCCAGGCGCTTCAGATG 291

Qy 181 GACATCTGAAAGCGGAGAGAGCCCTGCGGGAGCGAGGGCGCTGAAGCGCCGTCGAG 240

Db 292 GACATCTGAAAGCGGAGAGAGCCCTGCGGGAGCGAGGGCGCTGAAGCGCCGTCGAG 351

Qy 241 GGAGACATCATTTATCAGCGAGGAGCGGAGCTCTTCCGCGCCCTCCGCTCCAGGAC 300

Db 352 GGAGACATCATTTATCAGCGAGGAGCGGAGCTCTTCCGCGCCCTCCGCTCCAGGAC 411

Qy 301 CAGGTGGAGGTGGTGGCGAATTCGGGGGCGACGACAAAGCCCAAAATCAGGGGTGGAG 360

Db 412 CAGGTGGAGGTGGTGGCGAATTCGGGGGCGACGACAAAGCCCAAAATCAGGGGTGGAG 471

Qy 361 GAGGTGGAAAGCTGACACCGCATTCAGGGGAGTGTGCTGCTGGGCTTACCCCTAGCC 420

Db 472 GAGGTGGAAAGCTGACACCGCATTCAGGGGAGTGTGCTGCTGGGCTTACCCCTAGCC 531

Qy 421 ATCTCTCAGCGGCGCTTACCTGGGGTGTTCCTCATCATCTTCCGCGCGCTGCTGCTGC 480

Db 532 ATCTCTCAGCGGCGCTTACCTGGGGTGTTCCTCATCATCTTCCGCGCGCTGCTGCTGC 591

Qy 481 TACACCGCAAGATCCTCATTCGCGTGCCTGTACGAGGAGAAATGAAGACGCGAGGTGGTG 540

Db 592 TACACCGCAAGATCCTCATTCGCGTGCCTGTACGAGGAGAAATGAAGACGCGAGGTGGTG 651

Qy 541 CGCGTGGGAGCTCTGTAAGTGGCCATAGCCAAAGCTGTGCGCCCGCGCTTCCCAAG 600

Db 652 CGCGTGGGAGCTCTGTAAGTGGCCATAGCCAAAGCTGTGCGCCCGCGCTTCCCAAG 711

Qy 601 CTGGGCGCGAGTGGTGAACGTAGCGAGATCATCGAGCTGGTGTGATGAGTGCATCCTG 660

Db 712 CTGGGCGCGAGTGGTGAACGTAGCGAGATCATCGAGCTGGTGTGATGAGTGCATCCTG 771

Qy 661 TACGTGGTGGTGAAGTGGCAACCTCATGTACAAAGCTTCCCGGGGCTGCGCGCTGCGAG 720

Db 772 TACGTGGTGGTGAAGTGGCAACCTCATGTACAAAGCTTCCCGGGGCTGCGCGCTGCGAG 831

Qy 721 AGTCTCTGTCCATTTATCGCCAGCGCGCTGCTGCTGCTGCGCTTCCCTTAAGAACCTC 780

Db 832 AGTCTCTGTCCATTTATCGCCAGCGCGCTGCTGCTGCTGCGCTTCCCTTAAGAACCTC 891

Qy 781 AAGGCGCGTCCAAAGTTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840

Db 892 AAGGCGCGTCCAAAGTTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 951

Qy 841 GTCATAGCCTACTGTCTATCGGGGCGCGCATTCGGGCGCTGGGAGAGGTCAAAGTTCATC 900

Db 952 GTCATAGCCTACTGTCTATCGGGGCGCGCATTCGGGCGCTGGGAGAGGTCAAAGTTCATC 1011

Qy 901 ATCGAGTCAAGAGTTCCTCATTCCTTGGCATCATCGTGTTCAGTTCAGTTCAGTTCAG 960

Db 1012 ATCGAGTCAAGAGTTCCTCATTCCTTGGCATCATCGTGTTCAGTTCAGTTCAGTTCAG 1071

Qy 961 ATCTTCTGCTTCCGCTTCCGCTGGAGGCAATATGACAGACCCAGCGAGTTCACATCATGATG 1020



Db	1461	GTCTTACGCTGTCTCATGGCCATTATATGTGCGGCATCTTCGCGCTGTCTCATGGGCTCACC	1521
Qy	1381	GGCAGCCTCACGGGCGCGGCTCTGTGTTCTTGTGTGCCAGCCTCTTTACCTTCGCGCTG	1440
Db	1521	GGCAGCCTCACGGGCGCGGCTCTGTGTTCTTGTGTGCCAGCCTCTTTACCTTCGCGCTG	1580
Qy	1441	CTCTGGCGCAAGCTGTGTGGCACCAAGCTCTTCTTCGACGTGCGCATCTTCGTTCATCGGC	1500
Db	1581	CTCTGGCGCAAGCTGTGTGGCACCAAGCTCTTCTTCGACGTGCGCATCTTCGTTCATCGGC	1640
Qy	1501	GGCATCTGCAGCGTGTCCGGCTTGTGTGTCACCTCCCTCGAGGGCCTCATCGAAGCCTACCGA	1560
Db	1641	GGCATCTGCAGCGTGTCCGGCTTGTGTGTCACCTCCCTCGAGGGCCTCATCGAAGCCTACCGA	1700
Qy	1561	ACCAACGCGGAGGACTAG	1578
Db	1701	ACCAACGCGGAGGACTAG	1718
RESULT 7			
US-10-239-566-1			
; Sequence 1, Application US/10239566			
; Publication No. US20030069398A1			
; GENERAL INFORMATION:			
; APPLICANT: Merck Patent GmbH			
; TITLE OF INVENTION: New human GABA recetor			
; FILE REFERENCE: VGATFRWS			
; CURRENT APPLICATION NUMBER: US/10/239,566			
; CURRENT FILING DATE: 2002-09-24			
; NUMBER OF SEQ ID NOS: 2			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 1			
; LENGTH: 1800			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (100)..(1674)			
US-10-239-566-1			
Query March			
Best Local Similarity 99.9%; Score 1576.4; DB 14; Length 1800;			
Matches 1577; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
Qy	1	ATGGCCACCTTGTCTCGGCAGCAAGCTGTCCAAAGTGGCCACGTCCGTGTGCCAACAAAGTCC	60
Db	100	ATGGCCACCTTGTCTCGGCAGCAAGCTGTCCAAAGTGGCCACGTCCGTGTGCCAACAAAGTCC	159
Qy	61	CAGGCCAAGATGAGCGGCATGTTCCGCCAGATGGTTTTCAGGCGGCCACCGATGAGGAG	120
Db	160	CAGGCCAAGATGAGCGGCATGTTCCGCCAGATGGTTTTCAGGCGGCCACCGATGAGGAG	219
Qy	121	CGGTGGCTTCGGCATTTGCCACCTTCGACCTTTGAGCACCGCCAGGGCTCGCAGATG	180
Db	220	CGGTGGCTTCGGCATTTGCCACCTTCGACCTTTGAGCACCGCCAGGGCTCGCAGATG	279
Qy	181	GACATCTCTAAAGCCGAGAGAGCCCTCGCGGGACGAGGCGCTGAAGCGCCGCTCGAG	240
Db	280	GACATCTCTAAAGCCGAGAGAGCCCTCGCGGGACGAGGCGCTGAAGCGCCGCTCGAG	339
Qy	241	GGAGACATCCATTATCAGCGAGGACGCGAGTCTCTTCGCCCTCCGGTTCCTCAAGGAC	300
Db	340	GGAGACATCCATTATCAGCGAGGACGCGAGTCTCTCTCGCCCTCCGGTTCCTCAAGGAC	399
Qy	301	CAGTGGAGGTGTGGCGAATTCCGGGGCCACGACCAAGCCCAAAATCAGCGCTGGAG	360
Db	400	CAGTGGAGGTGTGGCGAATTCCGGGGCCACGACCAAGCCCAAAATCAGCGCTGGAG	459
Qy	361	CGAGCTGGAACGTGACCAACGCCATCCAGGCGATGTTCTGTCTGGGCTACCCCTACGCC	420
Db	460	CGAGCTGGAACGTGACCAACGCCATCCAGGCGATGTTCTGTCTGGGCTACCCCTACGCC	519
Qy	421	ATCTTGCAACGCGGCTACCTGGGGTGTGTTTCTCATCATCTTCTGCCGCGGTTGTGTGCTGC	480

```
; APPLICANT: Gunther, Erik
; APPLICANT: Millet, Isabelle
; APPLICANT: Stone, David J
; APPLICANT: Smithson, Glennnda
; APPLICANT: Szekeres Jr, Edward S
; APPLICANT: Ji, Weizhen
; TITLE OF INVENTION: No. US20040068095A1el Human Proteins, Polynucleotides Encoding Th
; FILE REFERENCE: 21402-141 CJP
; CURRENT APPLICATION NUMBER: US/10/096.625
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 09/972,211
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 60/238,325
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/238,323
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/238,400
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,397
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,401
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,379
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,402
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 30/238,384
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,373
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,372
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,383
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,382
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/275,892
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/296,860
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 200
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29
; LENGTH: 1763
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (141)..(1715)
US-10-096-625-29

Query Match          99.9%; Score 1576.4; DB 16; Length 1763;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1577; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCCACCTTGCTCCGACAGCTGTCCAAAGTGGCCACAGCTCGGTGTCCAAAGTCC 60
DB 141 ATGGCCACCTTGCTCCGACAGCTGTCCAAAGTGGCCACAGCTCGGTGTCCAAAGTCC 200
QY 61 CAGGCCAAGATGAGCGCATGTTCCGACGATGGTTTTCAGCGCGCCACGGATGAGGAG 120
DB 201 CAGGCCAAGATGAGCGCATGTTCCGACGATGGTTTTCAGCGCGCCACGGATGAGGAG 260
QY 121 CGCGTGGGCTTCGCGCATTCGACGACCTTCGACTTTGAGCACCGCCAGGCGCTTGCAGATG 180
DB 261 CGCGTGGGCTTCGCGCATTCGACGACCTTCGACTTTGAGCACCGCCAGGCGCTTGCAGATG 320
QY 181 GACATCTCTGAAGCCGAGGAGAGCCCTTCGGGGAGCGAGGCGCGCTGAGGCGCCCTCGAG 240
DB 321 GACATCTCTGAAGCCGAGGAGAGCCCTTCGGGGAGCGAGGCGCGCTGAGGCGCCCTCGAG 380
QY 241 GGAGACATCCATTATCAGCGAGGAGCGGAGCTCCTCTGCGCGCCCTCCGCGCTCCAAAGGAC 300

381 GGAGACATCCATTATCAGCGAGGAGCGGAGCTCCTCTGCGCGCCCTCCGCGCTCCAAAGGAC 440
301 CAGGTGGAGGTGGTGGCGGAATTCGGGGGCCACGACAAAGCCCAAAATACAGCGGTGGAG 360
441 CAGGTGGAGGTGGTGGCGGAATTCGGGGGCCACGACAAAGCCCAAAATACAGCGGTGGAG 500
361 GCAGGCTGGAACTGACCAACGCGCATCCAGGCGCATGTTCTGCTGGGCGCTTACCCCTAGCC 420
501 GCAGGCTGGAACTGACCAACGCGCATCCAGGCGCATGTTCTGCTGGGCGCTTACCCCTAGCC 560
421 ATCTGTCACGCGCGCTACCTGGGGTGTGTTTCTCATCATCTTTCGCGCGCTTGTGTGCTGC 480
561 ATCTGTCACGCGCGCTACCTGGGGTGTGTTTCTCATCATCTTTCGCGCGCTTGTGTGCTGC 620
481 TACACCGGCAAGATCTCTCATCGCGTGTCTGTAAGAGAAATGAAGACGCGAGGTGGT 540
621 TACACCGGCAAGATCTCTCATCGCGTGTCTGTAAGAGAAATGAAGACGCGAGGTGGT 680
541 CGCGTGGGAGCTCGTACGTTGGCCATAGCAACGCGCTGTGCGCGCGCGCTTCCCAAGC 600
681 CGCGTGGGAGCTCGTACGTTGGCCATAGCAACGCGCTGTGCGCGCGCGCTTCCCAAGC 740
601 CTGCGCGCGCGAGTGGTGAACGTAGCGCAGATCATCGAGCTGGTGTGATGAGTGCATCCTG 660
741 CTGCGCGCGCGAGTGGTGAACGTAGCGCAGATCATCGAGCTGGTGTGATGAGTGCATCCTG 800
661 TAGCTGGTGGTGGTGGCAACCTCATGTACAAAGCTTCCCGGGGTGCGCGTGTGCGAG 720
801 TAGCTGGTGGTGGTGGCAACCTCATGTACAAAGCTTCCCGGGGTGCGCGTGTGCGAG 860
721 AAGTCTGTGTCATTATCGCCACGCGCGTGTCTGCTGCTGCGCTTCCCTTAGAAGCTC 780
861 AAGTCTGTGTCATTATCGCCACGCGCGTGTCTGCTGCTGCGCTTCCCTTAGAAGCTC 920
781 AAGCGCGTGTCCAAAGTTCAGTCTGCTGTGCACCTCTGCGCCACCTTCGTCAATATCCTG 840
921 AAGCGCGTGTCCAAAGTTCAGTCTGCTGTGCACCTCTGCGCCACCTTCGTCAATATCCTG 980
841 GTCATAGCCTACTGTCTATCGCGGGCGCGAGCTGGGCGCTGGGAGAGGTCAGTGTCTAC 900
981 GTCATAGCCTACTGTCTATCGCGGGCGCGAGCTGGGCGCTGGGAGAGGTCAGTGTCTAC 1040
901 ATCGAGCTCAAGAAAGTTCCTCATCTGCGTGGCATCATCGTGTTCAGCTACAGTCTCAG 960
1041 ATCGAGCTCAAGAAAGTTCCTCATCTGCGTGGCATCATCGTGTTCAGCTACAGTCTCAG 1100
961 ATCTTCTGCTTCTGCTGGAGGCAATATGCAGCAGCCAGCAGGTTTCCACTGCATGATG 1020
1101 ATCTTCTGCTTCTGCTGGAGGCAATATGCAGCAGCCAGCAGGTTTCCACTGCATGATG 1160
1021 AACTGACGCACATCGCAGCTGCGTGTCTCAAGGGGCTCTTTCGCGTGTGCTGCGCTACCTC 1080
1161 AACTGACGCACATCGCAGCTGCGTGTCTCAAGGGGCTCTTTCGCGTGTGCTGCGCTACCTC 1220
1081 ACCTGGGCGCAGACGAGGAGGTCATCAGGATAACCTGCGCGGCTTCATCCCGGCC 1140
1221 ACCTGGGCGCAGACGAGGAGGTCATCAGGATAACCTGCGCGGCTTCATCCCGGCC 1280
1141 GTGGTCAACATCTTTCTGCTGGCCAAAGGCGCTGTTCTCTATCTCTGCGCATTTTGTGCC 1200
1281 GTGGTCAACATCTTTCTGCTGGCCAAAGGCGCTGTTCTCTATCTCTGCGCATTTTGTGCC 1340
1201 GCTGTGAGGTGCTGAGAAAGTGTCTTTTCCAGGAAGGAGCAGCGCGCTTTTTCGCGGCC 1260
1341 GCTGTGAGGTGCTGAGAAAGTGTCTTTTCCAGGAAGGAGCAGCGCGCTTTTTCGCGGCC 1400
1261 TGTACAGCGGCGACGCGGCGCTGAAGTCTTGGGGGCTGACGCTGCGCTGCGCGCTGCTC 1320
1401 TGTACAGCGGCGACGCGGCGCTGAAGTCTTGGGGGCTGACGCTGCGCTGCGCGCTGCTC 1460
1321 GTCTTACGCTGCTATGCGCCATTTTATGTGCGCGACTTTCGCGCTGTCTATGCGGCTCACC 1380
```

;; PRIOR FILING DATE: 2001-06-08  
;; NUMBER OF SEQ ID NOS: 198  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 29  
;; LENGTH: 1763  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: (141)..(1715)  
US-09-972-211-29

Query Match 99.9%; Score 1576.4; DB 11; Length 1763;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1577; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCCACCTTGTCTGCGAGCAAGCTGTCCAAAGTGGCCACGTCGCTGTCCAAAGTCC 60  
DB 141 ATGGCCACCTTGTCTGCGAGCAAGCTGTCCAAAGTGGCCACGTCGCTGTCCAAAGTCC 200

QY 61 CAGGCCAAGATGAGCGGCGATGTTCCGACGAGTGGTTTTCAGGCGGCCACGATGAGGAG 120  
DB 201 CAGGCCAAGATGAGCGGCGATGTTCCGACGAGTGGTTTTCAGGCGGCCACGATGAGGAG 260

QY 121 GCGGTGGGCTTCGCGCATTTGCGACGACCTTCGACTTTTGAACCGCGCGGCTTCGAGATG 180  
DB 261 GCGGTGGGCTTCGCGCATTTGCGACGACCTTCGACTTTTGAACCGCGCGGCTTCGAGATG 320

QY 181 GACATCTCTGAAGCCGAGGAGAGCCCTCGGGGACGAGGCGCTGAAGCGCCGCTCGAG 240  
DB 321 GACATCTCTGAAGCCGAGGAGAGCCCTCGGGGACGAGGCGCTGAAGCGCCGCTCGAG 380

QY 241 GGAGACATCATTTATCAGCGAGCGCGGAGCTCTCTGCGCGCTTCGCGCTTCAAGGAC 300  
DB 381 GGAGACATCATTTATCAGCGAGCGCGGAGCTCTCTGCGCGCTTCGCGCTTCAAGGAC 440

QY 301 CAGGTGGGAGTGGTGGCGAAATTCGGGGGCGCACGAAAGCCCAAAATCAAGCGCTGGGAG 360  
DB 441 CAGGTGGGAGTGGTGGCGAAATTCGGGGGCGCACGAAAGCCCAAAATCAAGCGCTGGGAG 500

QY 361 GCAGGTGGAACGTGACCAACGCGCATCCAGGGCATGTTCTGCTGGGCGCTTACCTACGCC 420  
DB 501 GCAGGTGGAACGTGACCAACGCGCATCCAGGGCATGTTCTGCTGGGCGCTTACCTACGCC 560

QY 421 ATCTCTCAGCGGCGTACCTGGGTTGTTCTCATCATCTTCGCGCGCTTGCTGCTGCTG 480  
DB 561 ATCTCTCAGCGGCGTACCTGGGTTGTTCTCATCATCTTCGCGCGCTTGCTGCTGCTG 620

QY 481 TACACCGGCAAGATCCTCATTCGCGTGCCTGTACGAGGAGATGAAGACGCGAGGTGGTG 540  
DB 621 TACACCGGCAAGATCCTCATTCGCGTGCCTGTACGAGGAGATGAAGACGCGAGGTGGTG 680

QY 541 CCGGTGCGGGAATCTGCTGAGTGGCCATAGCCGCTGTGCGCGCGCGCTTCCCAAG 600  
DB 681 CCGGTGCGGGAATCTGCTGAGTGGCCATAGCCGCTGTGCGCGCGCGCTTCCCAAG 740

QY 601 CTGGGCGCGGAGTGGTGAACGTAGCGAGATCATCGAGTGGTATGACGTGCATCTCG 660  
DB 741 CTGGGCGCGGAGTGGTGAACGTAGCGAGATCATCGAGTGGTATGACGTGCATCTCG 800

QY 661 TAGCTGTGTGAGTGGCAACCTCATGTACAGCTTCCCGGGGCTGCGCGCTGTCGCGAG 720  
DB 801 TAGCTGTGTGAGTGGCAACCTCATGTACAGCTTCCCGGGGCTGCGCGCTGTCGCGAG 860

QY 721 AAGTCTGTGTCATTATTCGCCACGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780  
DB 861 AAGTCTGTGTCATTATTCGCCACGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 920

QY 781 AAGGCGGTGTCAGTTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840  
DB 921 AAGGCGGTGTCAGTTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 980

QY 841 GTCATAGCCTACTGTCTATTCGCGGCGCGGCTGCGGCTGCGGAGAGGTCAAGTCTTCTAC 900

## RESULT 6

US-10-096-625-29  
; Sequence 29, Application US/10096625  
; Publication No. US20040068095A1  
; GENERAL INFORMATION:  
; APPLICANT: Shimkets, Richard A  
; APPLICANT: Taupier Jr, Raymond J  
; APPLICANT: Burgess, Catherine B  
; APPLICANT: Zerhusen, Bryan D  
; APPLICANT: Mezes, Peter S  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Malyankar, Uriel M  
; APPLICANT: Grose, William M  
; APPLICANT: Alsbrook II, John P  
; APPLICANT: Lepley, Denise M  
; APPLICANT: Spytek, Kimberly Ann  
; APPLICANT: Li, Li  
; APPLICANT: Edinger, Shlomit  
; APPLICANT: Gerlach, Valerie  
; APPLICANT: Ellerman, Karen  
; APPLICANT: MacDougall, John R

DB 981 GTCATAGCCTACTGTCTATCGCGCGCGGCTGGGCTGGGAGAGGTCAAGTCTTAC 1040

QY 901 ATCGACGTCAAGAAGTTCCCATCTCCATTGGCATCATCGTGTTCAGCTACACGTCTCAG 960

DB 1041 ATCGACGTCAAGAAGTTCCCATCTCCATTGGCATCATCGTGTTCAGCTACACGTCTCAG 1100

QY 961 ATCTTCTGCTCTGCTGAGGGCAATATGACAGAGCCAGCGAGTTCACCTGATGATG 1020

DB 1101 ATCTTCTGCTCTGCTGAGGGCAATATGACAGAGCCAGCGAGTTCACCTGATGATG 1160

QY 1021 AACTGGACGACATCGCAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080

DB 1161 AACTGGACGACATCGCAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1220

QY 1081 ACCTGGGCGGACGAGACCAAGGAGTTCATCAGGATACCTGCCGCGCTCATCCGCGCC 1140

DB 1221 ACCTGGGCGGACGAGACCAAGGAGTTCATCAGGATACCTGCCGCGCTCATCCGCGCC 1280

QY 1141 GTGGTCAACATCTTTTCTGCTGCGCAAGCGCTGTTGCTCTATCTCTGCTGCTGCTGCTG 1200

DB 1281 GTGGTCAACATCTTTTCTGCTGCGCAAGCGCTGTTGCTCTATCTCTGCTGCTGCTGCTG 1340

QY 1201 GCTGTCGAGGTGCTGGAGAGTCTCTTTCAGAGAGGAGCGCGGCTTTTTCGCGGCTG 1260

DB 1341 GCTGTCGAGGTGCTGGAGAGTCTCTTTCAGAGAGGAGCGCGGCTTTTTCGCGGCTG 1400

QY 1261 TGCTACAGCGGCGGAGCGGCGCTGAGGTCCTGGGGGCTGAGCGCTGCGCGCTCGTC 1320

DB 1401 TGCTACAGCGGCGGAGCGGCGCTGAGGTCCTGGGGGCTGAGCGCTGCGCGCTCGTC 1460

QY 1321 GTCTTACGCTGCTCATGCGCATTTATGTCGCGCATTTCTGCGGCTGCTCATGGGCTCAC 1380

DB 1461 GTCTTACGCTGCTCATGCGCATTTATGTCGCGCATTTCTGCGGCTGCTCATGGGCTCAC 1520

QY 1381 GCGACCTTACGCGGCGCGGCTCTGTTTCTGTCGCGGAGCTCTTTCACCTGCGGCTG 1440

DB 1521 GCGACCTTACGCGGCGCGGCTCTGTTTCTGTCGCGGAGCTCTTTCACCTGCGGCTG 1580

QY 1441 CTCTGGCGCAAGCTGCTGTCGACCAAGTCTTTCGACGTCGCGCATCTTCTGCTCATCGG 1500

DB 1581 CTCTGGCGCAAGCTGCTGTCGACCAAGTCTTTCGACGTCGCGCATCTTCTGCTCATCGG 1640

QY 1501 GGCATCTGACGCTGCTGCGGCTTCTGTCATCTCCCTCGAGGGGCTCATCGAAGCTACCGA 1560

DB 1641 GGCATCTGACGCTGCTGCGGCTTCTGTCATCTCCCTCGAGGGGCTCATCGAAGCTACCGA 1700

QY 1561 ACCAAGCGGAGGACTAG 1578

DB 1701 ACCAAGCGGAGGACTAG 1718

Db 548 CAGGTGGAGGTGTGGCAATTCGGGGCCACGACCAAGCCAAATATCACGGGTGGAG 607  
Qy 361 GCAGGCTGGAAAGTGAACAAGCCATCCAGGGCATGTTGTCGTGGGCCATCCCTAGCC 420  
Db 608 GCAGGCTGGAAAGTGAACAAGCCATCCAGGGCATGTTGTCGTGGGCCATCCCTAGCC 667  
Qy 421 ATCTGTGACGCGGCTACCTCTGGGTGTTCTTCTCATCATCTTCGCGCGCTTGTGTGTCG 480  
Db 668 ATCTGTGACGCGGCTACCTCTGGGTGTTCTTCTCATCATCTTCGCGCGCTTGTGTGTCG 727  
Qy 481 TACACCGCAAGATTCCTCATCGCGTGCCTGTACGAGGAGAATGAAGACGCGAGGTGGTG 540  
Db 728 TACACCGCAAGATTCCTCATCGCGTGCCTGTACGAGGAGAATGAAGACGCGAGGTGGTG 787  
Qy 541 CGCGTGGGACTCGTAGTGGGCATAGCCAAAGCCCTGCTGCGCCCGCGCTTCCCAAGC 600  
Db 788 CGCGTGGGACTCGTAGTGGGCATAGCCAAAGCCCTGCTGCGCCCGCGCTTCCCAAGC 847  
Qy 601 CTGGCGGCGAGTGGTGAAGTAGCGCAGATCATCGAGCTGGTGATGAGTGCATCCTG 660  
Db 848 CTGGCGGCGAGTGGTGAAGTAGCGCAGATCATCGAGCTGGTGATGAGTGCATCCTG 907  
Qy 661 TACGTGTGTGAGTGGCAACCTCATGTACAAAGCTTCCCGGGCTGCCCGTGTGCGAG 720  
Db 908 TACGTGTGTGAGTGGCAACCTCATGTACAAAGCTTCCCGGGCTGCCCGTGTGCGAG 967  
Qy 721 AAGTCTGTGTCATATTCGCAAGCGCGTGTGCTGTCACATCGAGCTGGTGATGAGTGCATC 780  
Db 968 AAGTCTGTGTCATATTCGCAAGCGCGTGTGCTGTCACATCGAGCTGGTGATGAGTGCATC 1027  
Qy 781 AAGCGCGTGCAGTTCAGTCTGCTGTGACATCTGCCACATCGTGCATCATATCCTG 840  
Db 1028 AAGCGCGTGCAGTTCAGTCTGCTGTGACATCTGCCACATCGTGCATCATATCCTG 1087  
Qy 841 GTCATAGCTACTGTCTATCCGGCGCGCGACTGGGCCCTGGGAGAGGTCAAAGTCTTAC 900  
Db 1088 GTCATAGCTACTGTCTATCCGGCGCGCGACTGGGCCCTGGGAGAGGTCAAAGTCTTAC 1147  
Qy 901 ATCGAGCTCAAGAGTTCCTCATCTCATTTGGCAATCATGTTTACAGTACAGTCTCAG 960  
Db 1148 ATCGAGCTCAAGAGTTCCTCATCTCATTTGGCAATCATGTTTACAGTACAGTCTCAG 1207  
Qy 961 ATCTTCTGCTCTGCTGGAGGCAATATGACAGCCAGCGAGTTCACACTGCATGATG 1020  
Db 1208 ATCTTCTGCTCTGCTGGAGGCAATATGACAGCCAGCGAGTTCACACTGCATGATG 1267  
Qy 1021 AACTGGACGCACATCGCAGCTGCTGCTCAAGGGCTCTTCGGGCTGCTGCGCTACCTC 1080  
Db 1268 AACTGGACGCACATCGCAGCTGCTGCTCAAGGGCTCTTCGGGCTGCTGCGCTACCTC 1327  
Qy 1081 ACCTGGCCCGACGAGACCAAGAGGTTCATACCGATAAATCTGCGGCTCCATCCGCGCC 1140  
Db 1328 ACCTGGCCCGACGAGACCAAGAGGTTCATACCGATAAATCTGCGGCTCCATCCGCGCC 1387  
Qy 1141 GTGCTCAACATCTTCTGCTGGCAAGGCGCTGTGCTTCTATCTTGCCTATCTTGTGCC 1200  
Db 1388 GTGCTCAACATCTTCTGCTGGCAAGGCGCTGTGCTTCTATCTTGCCTATCTTGTGCC 1447  
Qy 1201 GCTGTGAGGTGTGGAGAGTGCCTCTTCCAGAGGAGCGCGCTTTTTCCCGGCC 1260  
Db 1448 GCTGTGAGGTGTGGAGAGTGCCTCTTCCAGAGGAGCGCGCTTTTTCCCGGCC 1507  
Qy 1261 TGCTACGCGGCGACGCGGCGCTGAAGTCTGGGGGCTGACGCTGCGCTGCGCGCTCGTC 1320  
Db 1508 TGCTACGCGGCGACGCGGCGCTGAAGTCTGGGGGCTGACGCTGCGCTGCGCGCTCGTC 1567  
Qy 1321 GTCTTACGCTGCTCATGGCCATTTATGTGCCACATTCGCGCTGCTCATGGGCTCAC 1380  
Db 1568 GTCTTACGCTGCTCATGGCCATTTATGTGCCACATTCGCGCTGCTCATGGGCTCAC 1627  
Qy 1381 GCGAGCTCACGGCGCGCGGCTCTGTGTTCTTGTGCTGCCAGGCTCTTTTACCTGCGGCTG 1440

Db 1628 GGCAAGCTACGGGCGCGGCTCTGTTTCTGCTGCCAGCTCTTTTACCTGCGGCTG 1687  
Qy 1441 CTCTGGCGCAAGTGTGTGGACCAAGTCTTTTTCGACGTTCGCATCTTTGTCATCGGC 1500  
Db 1688 CTCTGGCGCAAGTGTGTGGACCAAGTCTTTTTCGACGTTCGCATCTTTGTCATCGGC 1747  
Qy 1501 GGCATCTGCAGCTGTCCGGCTTTCGTGCACTCCCTCGAGGCTCATCGAAGCTTACCGA 1560  
Db 1748 GGCATCTGCAGCTGTCCGGCTTTCGTGCACTCCCTCGAGGCTCATCGAAGCTTACCGA 1807  
Qy 1561 ACCAACGCGGAGGACTAG 1578  
Db 1808 ACCAACGCGGAGGACTAG 1825

RESULT 5

US-09-972-211-29  
; Sequence 29, Application US/09972211  
; Publication No. US20040048245A1  
; GENERAL INFORMATION:  
; APPLICANT: Shimkets, Richard A  
; APPLICANT: Taupier Jr, Raymond J  
; APPLICANT: Burgess, Catherine E  
; APPLICANT: Zerhusen, Bryan D  
; APPLICANT: Mezes, Peter S  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Malyankar, Uriel M  
; APPLICANT: Grosse, William M  
; APPLICANT: Alsobrook II, John P  
; APPLICANT: Lepley, Denise M  
; APPLICANT: Spytek, Kimberly Ann  
; APPLICANT: Li, Li  
; APPLICANT: Edinger, Shlomit  
; APPLICANT: Gerlach, Valerie  
; APPLICANT: Ellerman, Karen  
; APPLICANT: MacDougall, John R  
; APPLICANT: Gunther, Erik  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Stone, David J  
; APPLICANT: Smithson, Glennda  
; APPLICANT: Szekeres Jr, Edward S  
; TITLE OF INVENTION: No. US20040048245A1 Human Proteins, Polynucleotides Encoding T  
; TITLE OF INVENTION: Methods Of Using The Same  
; FILE REFERENCE: 21402-141  
; CURRENT APPLICATION NUMBER: US/09/972,211  
; CURRENT FILING DATE: 2001-10-05  
; PRIOR APPLICATION NUMBER: 60/238,325  
; PRIOR FILING DATE: 2000-10-05  
; PRIOR APPLICATION NUMBER: 60/238,323  
; PRIOR FILING DATE: 2000-10-05  
; PRIOR APPLICATION NUMBER: 60/238,400  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/238,397  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/238,401  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/238,379  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/238,402  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/238,384  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/238,373  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/238,372  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/238,383  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/238,382  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/275,892  
; PRIOR FILING DATE: 2001-03-14  
; PRIOR APPLICATION NUMBER: 60/296,860

```
QY 421 ATCTGCGCGGCTACTCGGGTGTCTCTCATCATCTTCGCGCGCGTGTGTGCTGC 480
Db 829 ATCTGCGCGGCTACTCGGGTGTCTCTCATCATCTTCGCGCGCGTGTGTGCTGC 888
QY 481 TACACCGGCAAGATCCTCATCGCTGCTGTACGAGGAGATGAAGACGGGAGGTGGTG 540
Db 889 TACACCGGCAAGATCCTCATCGCTGCTGTACGAGGAGATGAAGACGGGAGGTGGTG 948
QY 541 CCGCTGCGGACTGTGATCGTGCCATAGCCAAAGCGCTGTGCGCCCGCGGCTTCCCAACG 600
Db 949 CCGCTGCGGACTGTGATCGTGCCATAGCCAAAGCGCTGTGCGCCCGCGGCTTCCCAACG 1008
QY 601 CTGGCGCGCGAGTGTGAAGCTAGCGAGATCATCGAGCTGTGATGAGCTGATCTCTG 660
Db 1009 CTGGCGCGCGAGTGTGAAGCTAGCGAGATCATCGAGCTGTGATGAGCTGATCTCTG 1068
QY 661 TACGTGGTGTGAGTGCGCAACCTCATGTACAAAGCTTCCCGGGCTGCGCGTGTGCGAG 720
Db 1069 TACGTGGTGTGAGTGCGCAACCTCATGTACAAAGCTTCCCGGGCTGCGCGTGTGCGAG 1128
QY 721 AAGTCTGTGTCATATGCGCAAGCGCGTGTGCTGCTGCGGCTTCTTAAAGAACCTC 780
Db 1129 AAGTCTGTGTCATATGCGCAAGCGCGTGTGCTGCTGCGGCTTCTTAAAGAACCTC 1188
QY 781 AAGGCGGTGTCAAGTTCAGTCTGTGTCACCTCTGGCCCACTTCGTATCAATATCTCTG 840
Db 1189 AAGGCGGTGTCAAGTTCAGTCTGTGTCACCTCTGGCCCACTTCGTATCAATATCTCTG 1248
QY 841 GTCATAGCCTACTGTCTATCGGGGCGCGCGACTGCGGCTGGGAGAGGTCAAGTCTTAC 900
Db 1249 GTCATAGCCTACTGTCTATCGGGGCGCGCGACTGCGGCTGGGAGAGGTCAAGTCTTAC 1308
QY 901 ATGAGCTCAAGAGTTCCTCATCTCATTTGGATCATCTGTGTCAGCTACAGCTCTAG 960
Db 1309 ATGAGCTCAAGAGTTCCTCATCTCATTTGGATCATCTGTGTCAGCTACAGCTCTAG 1368
QY 961 ATCTTCTCGCTTCTGCTGAGGCGCAATATGAGAGCGCCAGCGAGTTCACATGCATGATG 1020
Db 1369 ATCTTCTCGCTTCTGCTGAGGCGCAATATGAGAGCGCCAGCGAGTTCACATGCATGATG 1428
QY 1021 AACTGAGCACATCGACGCTCGTGTCTCAAGGGCTCTTCGCGCTGTGCGCTACCTC 1080
Db 1429 AACTGAGCACATCGACGCTCGTGTCTCAAGGGCTCTTCGCGCTGTGCGCTACCTC 1488
QY 1081 ACTTGGGCGAGAGAGAGTTCATCAAGATTAACCTGCGCGGCTCGATCGCGGC 1140
Db 1489 ACTTGGGCGAGAGAGAGTTCATCAAGATTAACCTGCGCGGCTCGATCGCGGC 1548
QY 1141 GTGGTCAACATCTTCTGTGGCCAAAGCGCTGTGCTCTATCTCTGCACTTCTTTGCG 1200
Db 1549 GTGGTCAACATCTTCTGTGGCCAAAGCGCTGTGCTCTATCTCTGCACTTCTTTGCG 1608
QY 1201 GCTGTGAGGTGTGAGAGTGTGCTCTTCCAGGAGGAGCGCGGCTTTTCCGGGCT 1260
Db 1609 GCTGTGAGGTGTGAGAGTGTGCTCTTCCAGGAGGAGCGCGGCTTTTCCGGGCT 1668
QY 1261 TGCTACAGGCGGAGCGGCGCTGAAGTCTGTGGGGCTGACGCTGGGCTGGCGCTGCTC 1320
Db 1669 TGCTACAGGCGGAGCGGCGCTGAAGTCTGTGGGGCTGACGCTGGGCTGGCGCTGCTC 1728
QY 1321 GTCTTCAAGCTGTCTATGSCATTTATGTGCGCACTTCGCGTGTCTATGGGCTTCAAC 1380
Db 1729 GTCTTCAAGCTGTCTATGSCATTTATGTGCGCACTTCGCGTGTCTATGGGCTTCAAC 1788
QY 1381 GGAGGCTTCAAGGCGCGCGCTCTGTCTTCTGCTGCGGAGGCTCTTTCACCTGCGGCTG 1440
Db 1789 GGAGGCTTCAAGGCGCGCGCTCTGTCTTCTGCTGCGGAGGCTCTTTCACCTGCGGCTG 1848
QY 1441 CTCTGCGCAAGCTGTGTGGCAACCAAGTCTTCTCGAGTTCGCACTTCTGTGCTACGCG 1500
Db 1849 CTCTGCGCAAGCTGTGTGGCAACCAAGTCTTCTCGAGTTCGCACTTCTGTGCTACGCG 1908

RESULT 4
US-10-789-241-9
; Sequence 9, Application US/10789241
; Publication No. US20040180332A1
; GENERAL INFORMATION:
; APPLICANT: Powell, Douglas M.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: AIDS AND HIV-RELATED DISORDERS USING 9145, 1725, 311, 837,
; TITLE OF INVENTION: 58305, 156, 14175, 50352, 32678, 5560, 7240, 8855, 12396,
; TITLE OF INVENTION: 12397, 13644, 19938, 2077, 1735, 1786, 10220, 17822, 33945,
; TITLE OF INVENTION: 43748, 47161, 81982 OR 46777
; FILE REFERENCE: MPI03-041PIRNMNIM
; CURRENT APPLICATION NUMBER: US/10/789,241
; CURRENT FILING DATE: 2004-02-27
; PRIOR APPLICATION NUMBER: US 60/454,202
; PRIOR FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: US 60/456,326
; PRIOR FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: US 60/465,240
; PRIOR FILING DATE: 2003-04-24
; PRIOR APPLICATION NUMBER: US 60/475,233
; PRIOR FILING DATE: 2003-06-02
; PRIOR APPLICATION NUMBER: US 60/478,952
; PRIOR FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: US 60/487,836
; PRIOR FILING DATE: 2003-07-16
; PRIOR APPLICATION NUMBER: US 60/500,111
; PRIOR FILING DATE: 2003-09-04
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 2585
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (248)...(1825)
US-10-789-241-9

Query Match 100.0%; Score 1578; DB 17; Length 2585;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCCACCTTGTCTCGGAGCAAGCTTCCAAAGTGGCCACGTCGCTGTCCAAAGTCC 60
Db 248 ATGGCCACCTTGTCTCGGAGCAAGCTTCCAAAGTGGCCACGTCGCTGTCCAAAGTCC 307
QY 61 CAGGCCAAGATCAGCGGCGATGTTCCAGGATGGGTTTTCAGGCGGCCACCGATGAGGAG 120
Db 308 CAGGCCAAGATCAGCGGCGATGTTCCAGGATGGGTTTTCAGGCGGCCACCGATGAGGAG 367
QY 121 GCGGTGGGCTTCGCGCATTTGCGACGACCTTCGACTTTGAGCACCGGCGGCTGCGAGATG 180
Db 368 GCGGTGGGCTTCGCGCATTTGCGACGACCTTCGACTTTGAGCACCGGCGGCTGCGAGATG 427
QY 181 GACATCTGAAGCCGAGGAGAGCCCTTCGCGGAGAGGCGCTGAGAGCGGCTTCGAG 240
Db 428 GACATCTGAAGCCGAGGAGAGCCCTTCGCGGAGAGGCGCTGAGAGCGGCTTCGAG 487
QY 241 GGAGACATCTCATTTATCAGCGAGGAGCGAGCTCTCTGCGGCTTCGAGGCTTCAAGGAC 300
Db 488 GGAGACATCTCATTTATCAGCGAGGAGCGAGCTCTCTGCGGCTTCGAGGCTTCAAGGAC 547
QY 301 CAGGTGGGAGGTGGTGGCGAATTCGGGGGCCACGACAAAGCCCAAAATCATCGGCGTGGGAG 360
```

Db 254 GSAGACATCCATTATACAGCAGGAGCAGCGAGCTCTCTCGCGCCCTCCAGGAC 313  
QY 301 CAGGTGGAGGTGGTGGCGAATTCGGGGGCCACGACAGCCCAAAATCAAGGCGTGGAG 360  
Db 314 CAGGTGGAGGTGGTGGCGAATTCGGGGGCCACGACAGCCCAAAATCAAGGCGTGGAG 373  
QY 361 GCAGGTGGAACGTGACCAACGCCATCCAGGCGATGTTCTGTGGGSCCTACCCATAGCC 420  
Db 374 GCAGGTGGAACGTGACCAACGCCATCCAGGCGATGTTCTGTGGGSCCTACCCATAGCC 433  
QY 421 ATCTGCACCGCGCTACCTGGGGTGTGTTCTCATCATCTTCGCGCGCTGTGTGCTGC 480  
Db 434 ATCTGCACCGCGCTACCTGGGGTGTGTTCTCATCATCTTCGCGCGCTGTGTGCTGC 493  
QY 481 TACACGGCAAGATCTCATCGGCTGCTGTACGAGAGAAATGAAGCGCGAGGTGGT 540  
Db 494 TACACGGCAAGATCTCATCGGCTGCTGTACGAGAGAAATGAAGCGCGAGGTGGT 553  
QY 541 CGGTGCGGACTCGTACGTGGCCATAGCAACGCTGCTGCGCGCCGCTTCCCAACG 600  
Db 554 CGGTGCGGACTCGTACGTGGCCATAGCAACGCTGCTGCGCGCCGCTTCCCAACG 613  
QY 601 CTGGCGCGCGAGTGGTGAACGTAGCGCAGATCATCGAGCTGGTGTGATGAGTGCATCCTG 660  
Db 614 CTGGCGCGCGAGTGGTGAACGTAGCGCAGATCATCGAGCTGGTGTGATGAGTGCATCCTG 673  
QY 661 TAGCTGGTGTGAGTGGCAACCTCATGTACACAGCTTCCGGGGTGCCTGTGTCGAG 720  
Db 674 TAGCTGGTGTGAGTGGCAACCTCATGTACACAGCTTCCGGGGTGCCTGTGTCGAG 733  
QY 721 AAGTCTGTGTCATTATCGCCACGCGCTGCTGCTGCTGCGCTTCCCTTAAGAACCTC 780  
Db 734 AAGTCTGTGTCATTATCGCCACGCGCTGCTGCTGCTGCGCTTCCCTTAAGAACCTC 793  
QY 781 AAGCGCGTGTCAAGTTTCAGTCTGCTGTGCACCTCTGCGCCACCTTCGTCAATATCCTG 840  
Db 794 AAGCGCGTGTCAAGTTTCAGTCTGCTGTGCACCTCTGCGCCACCTTCGTCAATATCCTG 853  
QY 841 GTCATAGCTACTGTCTATCGCGGGCGCGGACTGGGCTGGAGAGGTCAAGTTCATC 900  
Db 854 GTCATAGCTACTGTCTATCGCGGGCGCGGACTGGGCTGGAGAGGTCAAGTTCATC 913  
QY 901 ATCGAGCTCAAGAGTTCCCATCTCCATTGGCATCATCGTGTTCAGTACAGCTCTCAG 960  
Db 914 ATCGAGCTCAAGAGTTCCCATCTCCATTGGCATCATCGTGTTCAGTACAGCTCTCAG 973  
QY 961 ATCTTCTGCTTTCGCTGGAGGGCAATATGACAGCCAGCGAGTTCACACTGCATGATG 1020  
Db 974 ATCTTCTGCTTTCGCTGGAGGGCAATATGACAGCCAGCGAGTTCACACTGCATGATG 1033  
QY 1021 AACTGACGACATCGAGCTCGTCTCAAGGGCTCTTTCGCGCTCGTGCCTACCTC 1080  
Db 1034 AACTGACGACATCGAGCTCGTCTCAAGGGCTCTTTCGCGCTCGTGCCTACCTC 1093  
QY 1081 ACCTGGCGCAGACGAGACCAAGAGGTTCATACCGATAAATCTGCGCGGCTCCATCCGGCC 1140  
Db 1094 ACCTGGCGCAGACGAGACCAAGAGGTTCATACCGATAAATCTGCGCGGCTCCATCCGGCC 1153  
QY 1141 GTGGTCAACATCTTTCGTGGGCAAGGGGCTGTGTCTTATCTCTGCGCATTTCTTTGCG 1200  
Db 1154 GTGGTCAACATCTTTCGTGGGCAAGGGGCTGTGTCTTATCTCTGCGCATTTCTTTGCG 1213  
QY 1201 GCTGTGAGGTGTGAGAGTGCCTCTTCAGGAAGGAGCGCGGCTTTTTCGGGCG 1260  
Db 1214 GCTGTGAGGTGTGAGAGTGCCTCTTCAGGAAGGAGCGCGGCTTTTTCGGGCG 1273  
QY 1261 TGCTACAGCGGCGACGGCGCTGAAGTCTGCGGGCTGACGCTGCGCTGCGGCTCCTC 1320  
Db 1274 TGCTACAGCGGCGACGGCGCTGAAGTCTGCGGGCTGACGCTGCGCTGCGGCTCCTC 1333  
QY 1321 GTCTTACGCTGTCTATGGCCATTTATGTGCGCACTTCGCGCTGTCTATGGGCTTCACC 1380

Db 1334 GTCTTACGCTGTCTCATGGCCATTTATGTGCGCACTTATGTGCGTGTCTATGGGCTCAC 1393  
QY 1381 GGCAGGCTTCAAGGCGCGGCTCTCTGTTCTTGTCTGCCAGGCTCTTTTCACTTGGGCTG 1440  
Db 1394 GGCAGGCTTCAAGGCGCGGCTCTCTGTTCTTGTCTGCCAGGCTCTTTTCACTTGGGCTG 1453  
QY 1441 CTCTGCGCAAGCTGTGTGGCACCAAGTCTTCTTCCAGCTGCCCATCTTCTGTCATCGGC 1500  
Db 1454 CTCTGCGCAAGCTGTGTGGCACCAAGTCTTCTTCCAGCTGCCCATCTTCTGTCATCGGC 1513  
QY 1501 GGCATCTGAGCGTGTTCGCGCTTCTGTCACCTCCCTCGAGGGCTCATCGAAGCTTACCGA 1560  
Db 1514 GGCATCTGAGCGTGTTCGCGCTTCTGTCACCTCCCTCGAGGGCTCATCGAAGCTTACCGA 1573  
QY 1561 ACCAACCGCGAGGACTAG 1578  
Db 1574 ACCAACCGCGAGGACTAG 1591  
  
RESULT 3  
US-09-940-919-3  
; Sequence 3, Application US/09940919  
; Patent No. US20020082390A1  
; GENERAL INFORMATION:  
; APPLICANT: Friddle, Carl Johan  
; APPLICANT: Gerhardt, Brenda  
; APPLICANT: Hu, Yi  
; TITLE OF INVENTION: No. US20020082390A1e1 Human GABA Transporter Protein and Polynuc  
; FILE REFERENCE: LEX-0228-USA  
; CURRENT APPLICATION NUMBER: US/09/940,919  
; PRIOR FILING DATE: 2001-08-28  
; PRIOR APPLICATION NUMBER: US 60/230,178  
; PRIOR FILING DATE: 2000-09-01  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 2261  
; TYPE: DNA  
; ORGANISM: homo sapiens  
US-09-940-919-3  
  
Query Match 100.0%; Score 1578; DB 9; Length 2261;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 ATGGCCACCTTGTCTCGCAGCAAGCTGTCCAACTGGCCACGCTCCGCTGTCCAACTGTC 60  
Db 409 ATGGCCACCTTGTCTCGCAGCAAGCTGTCCAACTGGCCACGCTCCGCTGTCCAACTGTC 468  
QY 61 CAGGCCAAGATGAGCGGCATGTTGCGCAGGATGGGTTTTTCAGCGGCGCCACGATGAGGAG 120  
Db 469 CAGGCCAAGATGAGCGGCATGTTGCGCAGGATGGGTTTTTCAGCGGCGCCACGATGAGGAG 528  
QY 121 GCGGTGGGCTTCGCGCATTTGCGACACCTCGACTTTGAGCAGCGCCAGGCGCTGCAAGTG 180  
Db 529 GCGGTGGGCTTCGCGCATTTGCGACACCTCGACTTTGAGCAGCGCCAGGCGCTGCAAGTG 588  
QY 181 GACATCTTGAAGCGGAGGAGAGCCCTGCGGGGACGAGGCGCTGAAGCGCCGCTCGAG 240  
Db 589 GACATCTTGAAGCGGAGGAGAGCCCTGCGGGGACGAGGCGCTGAAGCGCCGCTCGAG 648  
QY 241 GGAGACATCATTTATCAGCGAGGACGAGAGCTCTCTGCGGCTTCGCGCTTCCAGGAC 300  
Db 649 GGAGACATCATTTATCAGCGAGGACGAGAGCTCTCTGCGGCTTCGCGCTTCCAGGAC 708  
QY 301 CAGGTGGGAGGTGGTGGCAATTCGGGGCCACGACCAAGCCCAAATACACGCGTGGGAG 360  
Db 709 CAGGTGGGAGGTGGTGGCAATTCGGGGCCACGACCAAGCCCAAATACACGCGTGGGAG 768  
QY 361 CGAGGCTGSAAGTGAACCAACGCCATCCAGGGCATGTTTGTGCTGTGGGCTTACCTACGCC 420  
Db 769 CGAGGCTGSAAGTGAACCAACGCCATCCAGGGCATGTTTGTGCTGTGGGCTTACCTACGCC 828

```
QY 121 GCGGTGGGCTTCGCGCATTTGCGACGACCTCGACTTTGAGCACCGCGCGGCTGCGATG 180
DB 121 GCGGTGGGCTTCGCGCATTTGCGACGACCTCGACTTTGAGCACCGCGCGGCTGCGATG 180
QY 181 GACATCTGAAAGCCGAGGAGAGCCCTGCGGGAGAGAGGGGCGCTGAAAGCGCCGCTCGAG 240
DB 181 GACATCTGAAAGCCGAGGAGAGCCCTGCGGGAGAGAGGGGCGCTGAAAGCGCCGCTCGAG 240
QY 241 GGAGACATCCATTATCAGGAGGAGAGGAGCTCTCTGCGCCCTCGGGCTCCAAAGGAC 300
DB 241 GGAGACATCCATTATCAGGAGGAGAGGAGCTCTCTGCGCCCTCGGGCTCCAAAGGAC 300
QY 301 CAGGTGGGAGGTGGTGGGGAATTCGGGGGCGCACGACAAAGCCCAAAATCACGCGCGTGGGAG 360
DB 301 CAGGTGGGAGGTGGTGGGGAATTCGGGGGCGCACGACAAAGCCCAAAATCACGCGCGTGGGAG 360
QY 361 GCAGGCTGGAACGTGACCAACGCCATCCAGGGCATGTTGCTGCTGGGCTTACCCTACGCC 420
DB 361 GCAGGCTGGAACGTGACCAACGCCATCCAGGGCATGTTGCTGCTGGGCTTACCCTACGCC 420
QY 421 ATCTGTCACGGCGGCTACTCTGCGGTGTTCTCATCATCTTCGCGCGGTTGTGTGCTGC 480
DB 421 ATCTGTCACGGCGGCTACTCTGCGGTGTTCTCATCATCTTCGCGCGGTTGTGTGCTGC 480
QY 481 TACACCGGCAAGATCTCATCGCTGCTGTACGAGGAGATGAAGACGGCGAGGTGGTG 540
DB 481 TACACCGGCAAGATCTCATCGCTGCTGTACGAGGAGATGAAGACGGCGAGGTGGTG 540
QY 541 CCGGTGCGGACTCGTACGTGCCATAGCCAAAGCGCTGCTGCGCCCGGCTTCCCAACG 600
DB 541 CCGGTGCGGACTCGTACGTGCCATAGCCAAAGCGCTGCTGCGCCCGGCTTCCCAACG 600
QY 601 CTGGGGCGGAGTGGTGAACCTAGCGAGATCATCGAGCTGGTGCATGACGCTGATCTCTG 660
DB 601 CTGGGGCGGAGTGGTGAACCTAGCGAGATCATCGAGCTGGTGCATGACGCTGATCTCTG 660
QY 661 TACGTGCTGCTGAGTGGCAACCTCATGTACAAAGCTTCCCGGGCTGCGCGTGTGCGAG 720
DB 661 TACGTGCTGCTGAGTGGCAACCTCATGTACAAAGCTTCCCGGGCTGCGCGTGTGCGAG 720
QY 721 AAGTCTGCTCATTTATGCGCAAGCGCGTGTGCTGCTGCGCGCTTCTTAAAGAACCTC 780
DB 721 AAGTCTGCTCATTTATGCGCAAGCGCGTGTGCTGCTGCGCGCTTCTTAAAGAACCTC 780
QY 781 AAGGCGGTGCAAGTTCAGTCTGCTGCACTCTGGGCCACTTCTGTCATCAATATCTG 840
DB 781 AAGGCGGTGCAAGTTCAGTCTGCTGCACTCTGGGCCACTTCTGTCATCAATATCTG 840
QY 841 GTCATAGCCTACTGTCTATCGGGCGCGGAGTGGGCTGGGAGAGGTCAAGTTCATC 900
DB 841 GTCATAGCCTACTGTCTATCGGGCGCGGAGTGGGCTGGGAGAGGTCAAGTTCATC 900
QY 901 ATCGAGCTCAAGAAGTTCCTCATTTGCGATCATCGTGTTCAGCTTACAGTTCAG 960
DB 901 ATCGAGCTCAAGAAGTTCCTCATTTGCGATCATCGTGTTCAGCTTACAGTTCAG 960
QY 961 ATCTTCTGCTGCTGCGTGGGAGGATATGAGGAGGAGGAGGAGTTCACCTGATGATG 1020
DB 961 ATCTTCTGCTGCTGCGTGGGAGGATATGAGGAGGAGGAGGAGTTCACCTGATGATG 1020
QY 1021 AACTGAGCCACATCGAGCTGCTGCTCAAGGGCTCTTTCGCGCTGCTGCGCTACCTC 1080
DB 1021 AACTGAGCCACATCGAGCTGCTGCTCAAGGGCTCTTTCGCGCTGCTGCGCTACCTC 1080
QY 1081 ACTTGGGCGGAGAGGAGGTTCATCAAGGATAACTTCCCGGCTTCAATCCGCGCC 1140
DB 1081 ACTTGGGCGGAGAGGAGGTTCATCAAGGATAACTTCCCGGCTTCAATCCGCGCC 1140
QY 1141 GTGGTCAACATCTTCTGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
DB 1141 GTGGTCAACATCTTCTGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
QY 1201 GCTGTGAGGTGCTGAGAGTCTCTTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
```

```
DB 1201 GCTGTGAGGTGCTGAGAGTCTCTTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260
QY 1261 TGCTACAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
DB 1261 TGCTACAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
QY 1321 GTCTTACAGCTGCTCATGCGCATTTATGTCGCGCATTTGTCGCGCTCATGCGGCTCAC 1380
DB 1321 GTCTTACAGCTGCTCATGCGCATTTATGTCGCGCATTTGTCGCGCTCATGCGGCTCAC 1380
QY 1381 GGCAAGCTCACAGGCGGCGGCTCTGTTTCTGTCGCGGAGGAGGAGGAGGAGGAGGAGGAG 1440
DB 1381 GGCAAGCTCACAGGCGGCGGCTCTGTTTCTGTCGCGGAGGAGGAGGAGGAGGAGGAGGAG 1440
QY 1441 CTCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500
DB 1441 CTCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500
QY 1501 GGCAATCTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560
DB 1501 GGCAATCTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560
QY 1561 ACCAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1620
DB 1561 ACCAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1620
```

## RESULT 2

```
US-09-871-503-1
; Sequence 1, Application US/09871503
; Patent No. US20020076758A1
; GENERAL INFORMATION:
; APPLICANT: George Christian Teratappen
; APPLICANT: Cinzia Felicita Sala
; TITLE OF INVENTION: POLYPEPTIDE
; FILE REFERENCE: QG 1013
; CURRENT APPLICATION NUMBER: US/09/871,503
; CURRENT FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: GB 0013239.9
; PRIOR FILING DATE: 2000-06-01
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1595
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (14)..(1588)
US-09-871-503-1
```

```
Query Match 100.0%; Score 1578; DB 9; Length 1595;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCCACTTGTCTCCGAGCAAGCTGTCACCAAGTGGGCTGTCGCTGTCACCAAGTTC 60
DB 14 ATGGCCACTTGTCTCCGAGCAAGCTGTCACCAAGTGGGCTGTCGCTGTCACCAAGTTC 73
QY 61 CAGGCCAAGATGAGCGGATGTTCCGCAAGATGGGTTTTCAGGGGGCCACCGATGAGGAG 120
DB 74 CAGGCCAAGATGAGCGGATGTTCCGCAAGATGGGTTTTCAGGGGGCCACCGATGAGGAG 133
QY 121 GCGGTGGGCTTTCGCGCATTTGCGACCTTCAGCTTTGAGCAGGAGGAGGAGGAGGAGGAG 180
DB 134 GCGGTGGGCTTTCGCGCATTTGCGACCTTCAGCTTTGAGCAGGAGGAGGAGGAGGAGGAG 193
QY 181 GACATCTTGAAGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240
DB 194 GACATCTTGAAGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 253
QY 241 GGAGACATCTTATTCAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
```







```
; Sequence 20, Application US/09230371A
; Patent No. 6348586
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A
; APPLICANT: Russo, James J
; APPLICANT: Edelman, Isidore S
; APPLICANT: Moore, Patrick S
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS SEQUENCES AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 45185-G-PCT-US
; CURRENT APPLICATION NUMBER: US/09/230,371A
; CURRENT FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: PCT/US97/13346
; PRIOR FILING DATE: 1997-07-22
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 32207
; TYPE: DNA
; ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-230-371A-20

Query Match          3.7%; Score 58; DB 3; Length 32207;
Best Local Similarity 52.5%; Pred.No. 0.0011;
Matches 127; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

QY      1317 CGTGTGCTTTACAGCGTGCTCATGGCCATTATATGCGGCACCTTCGGCGTGCTCATGGGCCT 1376
DB      20703 CATGCCCCCGGCTCCTCGTCTCCTCTCTGTGCTGCTCTGTCATCCTCCTCGTCT 20762

QY      1377 CACGGGACGCTCACGGGGCGCGGCTCTGTGTTCTTGCTGCCAGGCTCTTTACACTGGCG 1436
DB      20763 CATTGTCTCTCATCATGTCATCTCTCGTCTCCTCTCCTCGTCTCCTCCTCGTCT 20822

QY      1437 CCTGTCTCGCGCAAGCTGCTGTGGCACCAAGTCTTCTTCGACGTCGCGCATCTTCGTCTAT 1496
DB      20823 CCTCTCTGTCTCTCTCCTCGTCATCTCTCTCGTCATCTCTCTCGTCATCTCTCTCGTCT 20882

QY      1497 CGGCGGCATCTCGAGCGGTGTCGGGTTTCGTGCACTCCCTCGAGGGGCTCATCGAAGCCTA 1556
DB      20883 CCTCTCTCGTCATCTCTCTCGTCATCTCTCTCGTCATCTCTCTCGTCATCTCTCTCGTCT 20942

QY      1557 CC 1558
DB      20943 CC 20944
```

Search completed: November 3, 2004, 22:25:45  
Job time : 166 secs





[illegible]

TITLE OF INVENTION: Immediate Early Protein From Kaposi's  
 TITLE OF INVENTION: Sarcoma-Associated Herpesvirus, DNA  
 TITLE OF INVENTION: Encoding Same And Uses Thereof  
 NUMBER OF SEQUENCES: 21  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Cooper & Dunham LLP  
 STREET: 1185 Avenue of the Americas  
 CITY: New York  
 STATE: New York  
 COUNTRY: U.S.A.  
 ZIP: 10036  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/728,323A  
 FILING DATE:  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:

[illegible]

; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 2057  
; LENGTH: 981  
; TYPE: DNA  
; ORGANISM: Drosophila melanogaster  
US-09-270-767-2057

Query Match 7.0%; Score 110.8; DB 4; Length 981;  
Best Local Similarity 53.8%; Pred. No. 2.7e-16;  
Matches 229; Conservative 0; Mismatches 197; Indels 0; Gaps 0;

QY 1138 GCGGTGGTCAACATCTTTCTGGTGGCCAGGCGCTTGTCTATCTCTGCAATCTTT 1197  
DB 2 GGCATGGTGAATCTTCTCTGGTGCATCAAGGCCCTTCTGAGTACCCCTTGCCCTACTAT 61

QY 1198 GCGGTGTGCGAGGTGTGGAGAAGTCTCTTCCAGGAAGGCGCGGCCCTTTTCCCG 1257  
DB 62 GTGCTGTGAGCTGCTGAAACGCAATCTTTTAGAGGTTCGCCAAGACCAATTTCCA 121

QY 1258 GCGTGTACAGGGGACGGGCGCTGAAGTCTCTGGGGCTGACGCTGGCTGGCGCTC 1317  
DB 122 ACCATCTGGAATCTGACGGGAACTGAAGGTTTGGGGTCTGGGCTTCGGTGTGGCGTC 181

QY 1318 GTGCTTTCACGCTGCTCATGCCCAATTTATGTGCGCACTTCGCGCTGCTCATGGGCTC 1377  
DB 182 ATTGTATCCACCATCTGATGCCAATTTTCATTCGCCACTTCTCCATCTGATGGGTTTC 241

QY 1378 ACCGGCAGCTCACGGGCGCGGCTCTGTTTCTGCTGCCAGGCTCTTTCACCTGGGC 1437  
DB 242 ATTGCGAGCTTCACTGGCAATGCTCAGCTTTAATCTGCGCTTGTATTTCCACATCAAG 301

QY 1438 CTGCTCTGGCGAAGCTGTGTGGACCAAGTCTTCTTCGAGTGGCCATCTTGTCTATC 1497  
DB 302 ATCAAGGGGCATCTGCTGATCAGAAGGAATAGCAAGACTACCTCATCTTGGGCTC 361

QY 1498 GCGGCATCTGACGGTGTCCGGCTCTGTCATCTCCCTCGAGGGCTCATCAAGCCTAC 1557  
DB 362 GCGTGTCTTCCGGCTTATGTGATCTACGATTCGGCAATGCGCTGATTAATGCAATTT 421

QY 1558 CGAACC 1563  
DB 422 GAAATC 427

RESULT 6  
US-09-270-767-17339  
; Sequence 17339, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 17339  
; LENGTH: 981  
; TYPE: DNA  
; ORGANISM: Drosophila melanogaster  
US-09-270-767-17339

Query Match 7.0%; Score 110.8; DB 4; Length 981;  
Best Local Similarity 53.8%; Pred. No. 2.7e-16;  
Matches 229; Conservative 0; Mismatches 197; Indels 0; Gaps 0;

QY 1138 GCGGTGGTCAACATCTTTCTGGTGGCCAGGCGCTGTTGTCTATCTCTGCAATCTTT 1197  
DB 2 GGCATGGTGAATCTTCTCTGGTGCATCAAGGCCCTTCTGAGCTACCCCTTGCCCTACTAT 61

QY 1198 GCCGCTGTGAGGTGCTGGAGAAGTCTCTTCCAGGAAGGACGCCGCGCTTTTCCCG 1257  
DB 62 GCTGCTGTGAGTGTCTGAAACGCAATCTTTTAGAGGTCCGCCAAGACCAATTTCCA 121

QY 1258 GCGTGTACAGGGGAGCGGCGCTGAAGTCTTGGGGCTGACGCTGGCTGGCGCTC 1317  
DB 122 ACCATCTGGAATCTGGAACGCGGAACTGAAGGTTTGGGCTTCCGCTGTCGGCGTC 181

QY 1318 GTGCTTTCACGCTGCTCATGGCCATTTATGTGCGCACTTCTGGGCTGCTCATGGGCTC 1377  
DB 182 ATTGTATCCACCATCTGATGGCCATTTTCAATCCCACTTCTCCATCTGATGGGTTTC 241

QY 1378 ACCGGCAGCTCACGGGCGCGGCTCTGTTTCTGCTGCCAGGCTCTTTCACCTGGGC 1437  
DB 242 ATTGCGAGCTTCACTGGCAATGCTCAGCTTTAATCTGCGCTTGTATTTCCACATCAAG 301

QY 1438 CTGCTCTGGCGAAGTCTGCTGSCACCAAGTCTTCTTCGAGCTGGCCATCTTCTGTCATC 1497  
DB 302 ATCAAGGGGCATCTGCTGATCAGAAGGAATAGCAAGACTACCTCATCTTGGGCTC 361

QY 1498 GCGGCATCTGACGGTGTCCGGCTTGTGCACTCCCTCGAGGGCTCATCAAGCCTAC 1557  
DB 362 GCGTGTCTTCCGGGTTATTTGTTATCTACGATTCGGCAATGCGCTGATTAATGCAATTT 421

QY 1558 CGAACC 1563  
DB 422 GAAATC 427

RESULT 7  
US-09-249-585A-4  
; Sequence 4, Application US/09249585A  
; Patent No. 6417002  
; GENERAL INFORMATION:  
; APPLICANT: Horlick, Robert  
; TITLE OF INVENTION: METHOD FOR MAINTENANCE AND SELECTION OF EPISOMES  
; FILE REFERENCE: 0867/0D905  
; CURRENT APPLICATION NUMBER: US/09/249,585A  
; CURRENT FILING DATE: 1999-02-11  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: Patentin version 3.0  
; SEQ ID NO 4  
; LENGTH: 1926  
; TYPE: DNA  
; ORGANISM: Epstein Barr Virus  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(1926)  
; OTHER INFORMATION: template strand of EBNA-1 DNA  
US-09-249-585A-4

Query Match 4.4%; Score 68.8; DB 4; Length 1926;  
Best Local Similarity 45.8%; Pred. No. 1.5e-06;  
Matches 238; Conservative 0; Mismatches 282; Indels 0; Gaps 0;

QY 1030 CACATCGCAGCTGCGTCTCAAGGGCCTTTCGGGCTCTGTCGCTTACCTACCTGGGC 1089  
DB 269 CTCGTCTCTGCTCTGCGCTCCCGCTCTGCTCTCTCCCGTCTCTGCTCTCTCCCGTC 328

QY 1090 GACGAGACCAAGAGGTATCAGGATAACCTGCCCGGCTTCCATCGGCGCGTGGTCAAC 1149  
DB 329 CTCGTCTCTCTCCCGTCTCTCCCGTCTCTCCCGTCTCTCTCTCTCTCTCTCTCC 388

QY 1150 ATCTTTCTGGTGGCCAAAGCGCTGTTGCTATCTCTGCGCATCTTTTTCGCGCTGTCGAG 1209  
DB 389 GTCTCTGCTCTCTCTCCCGTCTCTCCCGTCTCTCTCTCTCTCTCTCTCTCTCTCT 448

QY 1210 GTGCTGGAGAAGTCTCTTCCAGGAAGGACGCGCGCTTTTTCGCGGCTGTACAGC 1269  
DB 449 GTCTCTCTCTCCCGTCTCTCCCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 508

QY 1270 GCGCAGCGGCGCTGAAAGTCTTGGGGGTGACGCTGCGGTGCGCGCTGCTGCTCTTACG 1329

```
RESULT 3
US-09-270-767-15610/c
; Sequence 15610, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15610
; LENGTH: 448
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-15610
Query Match      8.8%; Score 139.4; DB 4; Length 448;
Best Local Similarity 60.0%; Pred. No. 5.6e-23;
Matches 274; Conservative 0; Mismatches 171; Indels 12; Gaps 2;

Qy 369 GAACGTGACCAACCCATCCAGGCATGTTGCTGCTGGGCTACCCCTACGCCATCCCTGCA 428
Db 448 GAACGTGACCAATGCCATTCAGGCATGTTTCATGTCGTGCGCTTGGCGCTCTCCA 389
Qy 429 CGGCGGCTACCTGGGGTGTTCATCATCTTCGCGCGCGTGTGCTGTGCTACACGG 488
Db 388 TGGCGGCTATTGGGCCATCGTGGCCATGGTGGGAATAGCACATCTGCTGTATACGGG 329
Qy 489 CAAGATCTCATCGCTGCTGTACGA---GGAGAAAGAACAGCGCGAGTGTGCGCGT 545
Db 328 CAAAGTCTGTGTCAGTGTCTATACGAACCGGATCCGCGACGGGTCAAATGTCGAGT 269
Qy 546 GCGGGACTCGTACGTGCGCCATAGCAACCGCTGTGCGCGCGCGCTTCCCAACGCTGGG 605
Db 268 GCGCGACAGTTATGCGGCCATAGCAAGGTATGTTTG-----GTCCAAAGTTGGG 218
Qy 606 CGGCGGAGTGTGACGATGAGGAGATGATCGAGCTGGTATGACGTGATCCTGTACGT 665
Db 217 CGCGCGGCGAGTCAGCATTCGCCAGCTTATCGAGCTTCTGATGACCTGCTCTGTACGT 158
Qy 666 GGTGTGAGTGGCAACCTCATGTACAACAGCTTCCCGGGCTGCCCGTGTGCGAGAATC 725
Db 157 GGTGTGTGCGGATCATTTGGCCGAAAGTATCCACAGGGTCTGTTCCACTCGCATC 98
Qy 726 CTGCTCCATTATCGCCACGCGGTGCTGTGCTGCTGCGCTTCCCTTAAGAACCTCAAGGC 785
Db 97 CTGGATGCTATTCTGGGGCATATTCTCTGCTGCGCATGGGATTCCTTAAGTCACTGAAAT 38
Qy 786 CGTGTCCAAGTTCAGTCTGCTGTGCACTCTGGGCCAC 822
Db 37 GGTATCAACGCTCTCGTTCTGGTGCAATGTGCGCAC 1

RESULT 4
US-09-178-093B-3
; Sequence 3, Application US/09178093B
; Patent No. 6660846
; GENERAL INFORMATION:
; APPLICANT: Robert H. Edwards
; APPLICANT: Richard J. Reimer
; APPLICANT: Steve L. McIntire
; APPLICANT: Erik M. Jorgenson
; APPLICANT: Kim Schuke
; TITLE OF INVENTION: Vesicular Amino Acid Transported
; FILE REFERENCE: Composition and Method
; CURRENT APPLICATION NUMBER: US/09/178,093B
; CURRENT FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 60/063,012
; PRIOR FILING DATE: 1997-10-23
```

```
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 1586
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
; NAME/KEY: misc feature
; LOCATION: (1)-(1586)
; OTHER INFORMATION: UNC-47 cDNA
US-09-178-093B-3
Query Match      8.7%; Score 137.2; DB 4; Length 1586;
Best Local Similarity 50.6%; Pred. No. 2.7e-22;
Matches 359; Conservative 0; Mismatches 348; Indels 3; Gaps 1;

Qy 727 TGGTCCATTATCGCCACGGCGGTGCTGCTGCGCTTGGCGCTTCTCTTAAGAACCTCAAGGCC 786
Db 610 TGGATGATGATTACCTCAGCATCTTTACTAACGTGCTCAATTTCTTGATGATCTACAATTT 669
Qy 787 GTCTCCAAGTTCACTGCTGCTGCATCTGCGCCACTTCTGTCATCAATATCTCTGGTCA 846
Db 670 GTGTCTCGTTGTTCATTTTCAATGCAATATCTCATTTGATTTGTCATCTGATCATGGTC 729
Qy 847 GCCTACTGTCTATCGCGGGCGCGACTGGGCTGGGAGAGGTCAAGTTCTATCATCGAC 906
Db 730 CTTTACTGTCTGATTCGTCTCACAATGCTCTTCTCAACGATCACATTTTCATTGAAT 789
Qy 907 GTCAGAAGTTCCCATCTCCATTGGCATCATCGTGTTCAGCTACAGCTCTCAGATCTTC 966
Db 790 ATCAACACTCTCCGCAATTTGGAAATGGTTGTTTTCGGCTACACATCTCATATATTC 849
Qy 967 CTCGCTTCGTGAGGCGCAATATGACAGCCAGCGAGTTCCACTGCATGATGAACCTGG 1026
Db 850 CTTTCAAATTTAGAAGGAAATATGAAAAATCTCTGCTCAATTCACAGTAATGTTTAAATGG 909
Qy 1027 ACCCATCATCGACCTGCGTGTCAAGGGCTCTTCGCGCTGCTGCGCTACCTCACCTGG 1086
Db 910 TCACACATCGCCCTGCTGTGTTCAAAGTTGTTTTGGAAATGCTCGGATTTCTCAATTT 969
Qy 1087 GCGCAGCAGACCAAGAGGTCAACAGGATAAAGTCTGCTTCCAGGAGGAGCGCGCTGTC 1143
Db 970 GGAGAGCTTACACAGAGGAAATTTCAAATTCCTGCTTAATCAATCATTTTAAATTC 1029
Qy 1144 GTCAACATCTTCTGTGTCGCAAGCGCTGTTGCTATCTCTGCTGCAATCTTTGCGCT 1203
Db 1030 GTGAACCTGATTTAGTGTGTCAGGCTCTTCTATCATATCCGTTGCCATTTCTATGAGCT 1089
Qy 1204 GTCGAGGTGCTGGAGAAGTCGCTTTCAGAGGAGGAGCGCGCTTTCCTCCGCGCTGC 1263
Db 1090 GTTCAACTTTTGAAGAACAAATTTGTTCTTGGATATCTTCAGACACATTTCAAGTTGT 1149
Qy 1264 TACAGCGGCGAGCGGCGCTGAAGTCTTGGGGCTGACGCTGCGCTGCGCGCTCGTCTG 1323
Db 1150 TATTACCGGATAAATCTTTACGTGAATGGCGGTACTTTTAAGAAATTTCTAGTCTT 1209
Qy 1324 TTCACGCTGCTATGCGCAATTTATGTCGCGACTTCCGCGTGTCTATGGGCTTCACCGGC 1383
Db 1210 TTCACACTTTTTCGTTGCAATATCAGTTCCATATTTGGTAGAGTTGATGGGATAGTTGA 1269
Qy 1384 AGCTTCACGCGCGCGCTGTTTCTGCTGCCCGAGCTCTTTTCACTT 1433
Db 1270 AATATTACAGGAACAATGTTTATCATTTATCTGCGCGGCACTATTCCACCT 1319

RESULT 5
US-09-270-767-2057
; Sequence 2057, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 3, 2004, 18:34:45 ; Search time 159 Seconds  
(without alignments)  
7054.244 Million cell updates/sec

Title: US-09-940-919-1  
Perfect score: 1578  
Sequence: 1 atggccacctgtctcgag.....gaaccaacggaggactag 1578

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 35539441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA: \*  
1: /cgn2\_6/ptodata/1/ina/5A COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1352.2	85.7	2392	4	US-09-178-093B-4
2	139.4	8.8	448	4	US-09-270-767-328
3	139.4	8.8	448	4	US-09-270-767-15610
4	137.2	8.7	1586	4	US-09-178-093B-3
5	110.8	7.0	981	4	US-09-270-767-2057
6	110.8	7.0	981	4	US-09-270-767-17339
7	68.8	4.4	1926	4	US-09-249-585A-4
8	68.8	4.4	1931	2	US-09-130-114-2
9	58	3.7	3489	3	US-08-728-323A-1
10	58	3.7	3489	3	US-09-298-568-1
11	58	3.7	3489	4	US-09-410-399-1
12	58	3.7	3489	4	US-09-894-379-1
13	58	3.7	32207	2	US-08-770-379-20
14	58	3.7	32207	3	US-08-757-669A-20
15	58	3.7	32207	3	US-09-230-371A-20
16	55.4	3.5	885	4	US-09-252-991A-2480
17	55.4	3.5	921	4	US-09-252-991A-2561
18	55.4	3.5	2532	4	US-09-252-991A-2356
19	53.8	3.4	1428	4	US-09-252-991A-14810
20	53.8	3.4	1428	4	US-09-252-991A-14802
21	53.8	3.4	1668	4	US-09-252-991A-14950
22	53.8	3.4	1926	4	US-09-249-585A-2
23	53.8	3.4	1926	4	US-09-410-399-3
24	53.8	3.4	2580	3	US-09-050-863-2
25	53.8	3.4	2580	3	US-09-359-081-2
26	53.8	3.4	5452	2	US-09-130-114-1
27	53.8	3.4	8705	4	US-09-647-344A-14

C 28	53.8	3.4	9600	3	US-08-910-647-1	Sequence 1, Appli
C 29	53.8	3.4	9600	4	US-09-620-925-1	Sequence 1, Appli
C 30	53.8	3.4	10596	1	US-07-884-811-15	Sequence 15, Appl
C 31	53.8	3.4	10596	1	US-07-885-971-15	Sequence 15, Appl
C 32	53.8	3.4	10596	1	US-08-087-783A-15	Sequence 15, Appl
C 33	53.8	3.4	10596	1	US-08-194-088B-15	Sequence 15, Appl
C 34	53.8	3.4	10596	2	US-08-194-087-15	Sequence 15, Appl
C 35	53.8	3.4	10596	5	PCT-US93-04648-15	Sequence 15, Appl
C 36	53.8	3.4	16080	4	US-09-724-566A-48	Sequence 48, Appl
C 37	52.8	3.3	2571	4	US-09-252-991A-9022	Sequence 9022, Ap
C 38	52.8	3.3	3150	4	US-09-252-991A-8765	Sequence 8765, Ap
C 39	52.8	3.3	3450	4	US-09-252-991A-8986	Sequence 8986, Ap
C 40	52.6	3.3	7218	1	US-08-232-463-14	Sequence 14, Appl
C 41	50.8	3.2	1578	4	US-09-252-991A-10972	Sequence 10972, A
C 42	50.8	3.2	2547	4	US-09-252-991A-11271	Sequence 11271, A
C 43	50.8	3.2	2964	4	US-09-252-991A-10912	Sequence 10912, A
C 44	50.2	3.2	4403765	3	US-09-103-840A-2	Sequence 2, Appli
C 45	50.2	3.2	4411529	3	US-09-103-840A-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-09-178-093B-4  
; Sequence 4, Application US/09178093B  
; Patent No. 6660846  
; GENERAL INFORMATION:  
; APPLICANT: Robert H. Edwards  
; APPLICANT: Richard J. Reimer  
; APPLICANT: Steve L. McIntire  
; APPLICANT: Erik M. Jorgenson  
; APPLICANT: Kim Schuske  
; TITLE OF INVENTION: Vesicular Amino Acid Transported  
; FILE OF INVENTION: Composition and Method  
; FILE REFERENCE: 2002-0005.30  
; CURRENT APPLICATION NUMBER: US/09/178,093B  
; CURRENT FILING DATE: 2001-08-20  
; PRIOR APPLICATION NUMBER: 60/063,012  
; PRIOR FILING DATE: 1997-10-23  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 4  
; LENGTH: 2392  
; TYPE: DNA  
; ORGANISM: Rattus norvegicus  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)..(2392)  
; OTHER INFORMATION: RUNC-47 cDNA  
US-09-178-093B-4

Query Match	85.7%	Score	1352.2	DB	4	Length	2392
Best Local Similarity	90.8%	Pred. No.	2.1e-301				
Matches	1433	Conservative	0	Mismatches	145	Indels	0
		Gaps	0				
QY	1	ATGCCACCTTGTCTCGCAGCAGCTGTCCACGTCGCCACGTCCTCGTGTCCACAGTCC	60				
DB	87	ATGCCACCTTGTCTCGCAGCAGCTGTCCACGTCGCCACGTCCTCGTGTCCACAGTCC	146				
QY	61	CAGGCCAAGATGAGCGGCATGTTCCGCGAGGATGGTTCAGCGGCCACGGATGAGGAG	120				
DB	147	CAGGCCAAGTGTAGCGGCATGTTCCGCGAGGATGGTTCAGCGGCCACGGATGAGGAG	206				
QY	121	CGCGTGGGCTTCGCGCATTCGACACCTTCGATTTAGCACCGCCAGGCGCTTCGAGATG	180				
DB	207	CGCGTGGGCTTCGCGCATTCGACACCTTCGATTTAGCACCGCCAGGCGCTTCGAGATG	266				
QY	181	GACATCTCTAAAGCCGAGGAGGCCCTCGGGGACGAGGCGCTGAAGGCCCGCTCGAG	240				
DB	267	GACATCTCTAAAGCCGAGGAGGCCCTCGGGGACGAGGCGCTGAAGGCCCGCTCGAG	326				
QY	241	GGAGACATCCATTATCAGCGAGGCGAGCTCTCTTCGCGGCCCTCCGCTCCAAGGAC	300				









```
XX AAS81399;
XX 13-FEB-2002 (first entry)
XX DNA encoding novel human diagnostic protein #17203.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US008631.
XX 31-MAR-2000; 2000US-00540217.
XX 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX P-PSDB; ABG17212.
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX Claim 1; SEQ ID NO 17203; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (II) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activities. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
XX coding sequences of the invention. Note: The sequence data for this
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 1095 BP; 240 A; 341 C; 338 G; 176 T; 0 U; 0 Other;
XX
XX Query Match 24.5%; Score 387.2; DB 5; Length 1095;
XX Best Local Similarity 99.2%; Pred. No. 2.9e-69;
XX Matches 389; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
XX 1 ATGCCACCTTGCTCCGAGCAAGCTGTCCAAAGTCCAGTCCGCTCCAAAGTCC 60
XX
XX 220 ATGCCACCTTGCTCCGAGCAAGCTGTCCAAAGTCCAGTCCGCTCCAAAGTCC 279
XX
XX 61 CAGCCCAAGATGAGCGCATGTTCCAGCATGGTTTTCAGGGCGCAGGATGAGG 120
XX
XX 280 CAGCCCAAGATGAGCGCATGTTCCAGCATGGTTTTCAGGGCGCAGGATGAGG 339
XX
XX 121 GCGGTGGGCTTCGCGCATTCGAGACCTCGACTTTTGAGCACCGCCAGGGCCTGCAGATG 180
XX
XX 340 GCGGTGGGCTTCGCGCATTCGAGACCTCGACTTTTGAGCACCGCCAGGGCCTGCAGATG 399
XX
XX 181 GACATCTGTAAGCCGAGGAGAGCCCTCGGGGACGAGGGCGTGAAGGCCGCTGAG 240
XX
XX 400 GACATCTGTAAGCCGAGGAGAGCCCTCGGGGACGAGGGCGTGAAGGCCGCTGAG 459
XX
XX 241 GGAGACATCATTTATCAGCGAGGAGGAGGAGTCTCTGCGGCCCTCCGGCTCCAAGGAC 300
XX
XX 460 GGAGACATCATTTATCAGCGAGGAGGAGGAGTCTCTGCGGCCCTCCGGCTCCAAGGAC 519
XX
XX 301 CAGGTGGGAGTGGTGGCGAATTCGGGGGCCACGACAAAGCCCAAAATCAGGCGTGGGAG 360
XX
XX 520 CAGGTGGGAGTGGTGGCGAATTCGGGGGCCACGACAAAGCCCAAAATCAGGCGTGGGAG 579
XX
XX 361 GCAGGCTGGAAACGTGACCAACGCCATCCAGGG 392
XX
XX 580 GCAGGCTGGAAACGTGACCAACGCCATCCAGTG 611
XX
XX
XX RESULT 14
XX ABL14045
XX ID ABL14045 standard; cDNA; 1650 BP.
XX AC ABL14045;
XX XX
XX DT 26-MAR-2002 (first entry)
XX XX
XX DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 36617.
XX XX
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical; gene; ss.
XX XX
XX OS Drosophila melanogaster.
XX XX
XX PN WO200171042-A2.
XX XX
XX PD 27-SEP-2001.
XX XX
XX PF 23-MAR-2001; 2001WO-US009231.
XX XX
XX PR 23-MAR-2000; 2000US-0191637P.
XX PR 11-JUL-2000; 2000US-00614150.
XX XX
XX PA (PEKE ) PE CORP NY.
XX XX
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX XX
XX DR WPI; 2001-656860/75.
XX DR P-PSDB; ABB69942.
XX XX
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signaling and cell-cell
XX PT interactions.
XX XX
XX PS Claim 1; SEQ ID NO 36617; 21pp + Sequence Listing; English.
XX XX
XX CC The invention relates to an isolated nucleic acid detection reagent
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is
XX CC useful in developmental biology and in elucidating cell signalling and
XX CC cell-cell interactions in higher eukaryotes for the development of
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention
XX CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
XX CC ABB72072). The sequence data for this patent did not form part of the
XX CC printed specification, but was obtained in electronic format directly
XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX XX
XX SQ Sequence 1650 BP; 375 A; 438 C; 445 G; 392 T; 0 U; 0 Other;
XX
XX Query Match 23.2%; Score 366; DB 4; Length 1650;
XX Best Local Similarity 57.6%; Pred. No. 6.4e-65;
XX Matches 720; Conservative 0; Mismatches 515; Indels 15; Gaps 3;
```

XX Mus sp.  
OS Synthetic.  
XX WO200136632-A2.  
XX 25-MAY-2001.  
XX 17-NOV-2000; 2000WO-IL000766.  
XX 17-NOV-1999; 99IL-00132978.  
XX 10-DEC-1999; 99IL-00133455.  
XX (COMP-) COMPUGEN LTD.  
XX Levine Z, David A, Azar I, Khosravi R, Bernstein J;  
XX WPI; 2001-336004/35.  
XX P-PSDB; AAU02987.  
XX Novel alternative splicing variants e.g. variant of angiotensin  
PT converting enzyme (ACEV), useful in identifying candidate compounds  
PT capable of binding to the variant and to detect anti-variant antibodies.  
XX  
XX Claim 1; Page 374; 519pp; English.  
XX The sequence represents a DNA encoding an angiotensin converting enzyme  
CC splice variant (ACEV) polypeptide. The polypeptides of the invention  
CC include variants of granulocyte colony stimulating factor receptor,  
CC glucagon, interleukin 6, platelet-derived endothelial cell growth factor,  
CC cyclin-dependent kinase inhibitor 1C, cellular tumour antigen p53, and  
CC vasoactive intestinal polypeptide receptor 2. The polypeptides and their  
CC associated nucleic acids are useful for identification of variant  
CC sequences and detection of candidate compounds capable of binding the  
CC molecules. The sequences of the invention can be used in the treatment  
CC and diagnosis of various disorders including cardiovascular diseases such  
CC as arteriosclerosis, myocardial infarction and coronary arterial  
CC thrombosis, renal diseases such as diabetic nephropathy, muscular  
CC diseases such as hypertrophy, immune disorders such as immune complex  
CC nephritis, multiple sclerosis, cancer, sarcoidosis, nonaroidotic  
CC pulmonary granulomatous diseases such as asbestosis and vascular  
CC pathologies involving an endothelial abnormality such as deep vein  
CC thrombosis  
XX  
XX Sequence 1723 BP; 364 A; 506 C; 534 G; 319 T; 0 U; 0 Other;  
Query Match 72.2%; Score 1140; DB 5; Length 1723;  
Best Local Similarity 91.4%; Pred. NO. 2.1e-222;  
Matches 1232; Conservative 0; Mismatches 110; Indels 6; Gaps 2;  
QY 1 ATGGCCACCTTGTCTCCGACGAGCTGTCCACGTCGGCCACCTCGGTGTCACCAAGTCC 60  
DB 1517 ATGGCCACCTTGTCTCCGACGAGCTGTCCACGTCGGCCACCTCGGTGTCACCAAGTCC 1458  
QY 61 CAGGCCAAGATGAGCGGCATGTTTCGCGAGGATGGGTTTTCAGGCGGCCACCGATGAGGAG 120  
DB 1457 CAGGCCAAGTGAAGCGGCATGTTTCGCGAGGATGGGTTTTCAGGCGGCCACCGATGAGGAA 1398  
QY 121 GCGGTGGCTTCGCGATTCGACGACCTCGACTTTGAGCAGCCGCGGCGCTCGCAGATG 180  
DB 1397 GCGGTGGCTTCGCGACCTCGACGATCTCGACTTTGAGCAGCCGCGGCGCTCGCAGATG 1338  
QY 181 GACATCTCAAGCCGAGGAGGAGCCCTCGGGGACGAGGCGCTGAAGCGCCGCTCGAG 240  
DB 1337 GACATCTCAAGTGAAGCGGCAGGAGCCCTCGGGGACGAGGCGCGAGAACTCCCGTCGAG 1278  
QY 241 GGAGACATCATTTATCAGGAGGAGGAGGAGTCTCTCTCGCGCCCTCCCGGCTCCAGGAC 300  
DB 1277 GGAGACATTCATTTATCAGC---GCGGCGGCGCTCTCTCGCACCTCTGCTCCAGGAC 1221  
QY 301 CAG---GTGGGAGGTGGTGGCGAATTTCCGGGGGCGCACGACAAAGCCCAAAATCAAGCGGTGG 357  
DB 1220 CAGGCGGTGGAGTGTGGGAGTTCCGGGGGTTCAGCAAAACCCCAAGATCAAGCGGTGG 1161

QY 358 GAGGAGGCTGGAACGTCGACCAAGCCATCAGGGCATGTTCTGCTGGGCTTACCCTAC 417  
DB 1160 GAAGCGGCTGGAACGTCGACCAATCAGGGCATGTTCTGCTGGGCTTACCCTAC 1101  
QY 418 GCCATCTTCACGCGGCTACCTCGGGTGTCTTCATCATCTTCGCGCGGTGTGTGC 477  
DB 1100 GCCATCTTCACGCGGCTACCTCGGGTGTCTTCATCATCTTCGCGCGGTGTGTGC 1041  
QY 478 TGCTACACCGGCAAGATCTCATCGCGTCCTGACGAGGAGATGAAGAGCGGCGGTG 537  
DB 1040 TGCTACACCGGCAAGATCTCATCGCGTCCTGACGAGGAGATGAAGAGCGGCGGTG 981  
QY 538 GTGCGGTGCGGGACTCGTACGTGCGCATAGCAACGCTGCTGCGCGCCCGCGCTTCCCA 597  
DB 980 GTGCGGTGCGGGACTCGTATGTGCGCATAGCAACGCTGCTGCGCGCTTCCCA 921  
QY 598 ACGTGGGCGCGGAGTGTGAACGTAGCGAGATCATCGAGCTGGTGAATGACGTGCATC 657  
DB 920 ACCCTGGGCGCGGCTGTCAATGTGCGCGAGATCATCGAGCTGGTGAATGACGTGCATC 861  
QY 658 CTGTACGTGGTGTGAGTGGCAACCTCATGTACACAGCTTCCCGGGGCTGCCCGTGTGC 717  
DB 860 TTGTACGTGGTGTGAGCGGCAACCTCATGTACACAGTTCCTCCCGGGGCTGCCCGTGTGC 801  
QY 718 CAGAACTCTGCTCCATTATCGCCACGCGCTGCTGCTGCTGCGCTTCTTAAAGAAC 777  
DB 800 CAGAACTCTGCTCCATTATCGCCACGCGCTGCTGCTGCGCTTCTTAAAGAAC 741  
QY 778 CTCAGGCGGTGTCCAAAGTTCAGTCTGTGTGCTGCTGCTGCGGCTTCTGCTCATCAATATC 837  
DB 740 CTCAGGCGGTGTCCAAAGTTCAGTCTGTGTGCTGCTGCTGCGGCTTCTGCTCATCAATATC 681  
QY 838 CTGCTCATAGCTTACTGCTATCGCGGCGCGGCTGCTGCTGCGGCTTCTGCTCATCAATATC 897  
DB 680 CTGCTCATAGCTTACTGCTATCGCGGCGCGGCTGCTGCTGCGGCTTCTGCTCATCAATATC 621  
QY 898 TACATCGACGTCAAGAAAGTTCCTCATCTCCATTTGCGCATCATCGTGTTCAGCTTACAGTCT 957  
DB 620 TACATCGACGTCAAGAAAGTTCCTCATCTCCATTTGCGCATCATCGTGTTCAGCTTACAGTCT 561  
QY 958 CAGATCTTCTGCTGCTGAGGCGCAATATGACGAGCCAGCGAGTTCACATGTCATG 1017  
DB 560 CAGATCTTCTGCTGCTGCTGAGGCGCAATATGACGAGCCAGCGAGTTCACATGTCATG 501  
QY 1018 ATGAACTGGACGACATCGACCTGCTGCTCAGGGGCTTCTTCCGCTGCTGCGCTTAC 1077  
DB 500 ATGAACTGGACGACATCGACCTGCTGCTCAGGGGCTTCTTCCGCTGCTGCGCTTAC 441  
QY 1078 CTCACCTGGGCGGACGAGACCAAGAGGTTCATCAGGATTAACCTGCGGCTTCCATCCGC 1137  
DB 440 CTCACCTGGGCGGACGAGACCAAGAGGTTCATCAGGATTAACCTGCGGCTTCCATCCGC 381  
QY 1138 GCGGTGTCAACATCTTCTGCTGCGCAAGGCGCTGTGTCTATCTTCTGCGCATTTT 1197  
DB 380 GCGGTGTCAACATCTTCTGCTGCGCAAGGCGCTGTGTCTATCTTCTGCGCATTTT 321  
QY 1198 GCGGTGTGAGGTGCTGAGAGTCTGCTTTCAGGAGGCGAGCGGCGCTTTCGCG 1257  
DB 320 GCGGTGTGAGGTGCTGAGAGTCTGCTTTCAGGAGGCGAGCGGCGCTTTCGCG 261  
QY 1258 GCTCTCTACGCGGCGGCGGCTGAAAGTCTTGGGGGCTGACGCTGCGCTGCGGCTC 1317  
DB 260 GCTCTCTACGCGGCGGCGGCTGAAAGTCTTGGGGGCTGACGCTGCGCTGCGGCTC 201  
QY 1318 GTGCTTTCAGCTGCTCATGCGCAATTT 1345  
DB 200 GTGCTTTCAGCTGCTCATGCGCAATCT 173

RESULT 13  
AAS81399  
ID AAS81399 standard; cdNA; 1095 BP.

CC sequences and detection of candidate compounds capable of binding the  
CC molecules. The sequences of the invention can be used in the treatment  
CC and diagnosis of various disorders including cardiovascular diseases such  
CC as arteriosclerosis, myocardial infarction and coronary arterial  
CC thrombosis, renal diseases such as diabetic nephropathy, muscular  
CC diseases such as hypertrophy, immune disorders such as immune complex  
CC nephritis, multiple sclerosis, cancer, sarcoidosis, nonarctoidotic  
CC pulmonary granulomatous diseases such as asbestosis and vascular  
CC pathologies involving an endothelial abnormality such as deep vein  
CC thrombosis  
XX  
SQ

Sequence 1648 BP; 278 A; 540 C; 480 G; 350 T; 0 U; 0 Other;

Query Match 85.2%; Score 1344.2; DB 5; Length 1648;  
Best Local Similarity 90.5%; Pred. No. 5.8e-264;  
Matches 1428; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

Qy 1 ATGGCCACCTTGTCTCCGAGCAAGCTGTCCAACTGGCCAGCTCCGCTGTCCAAAGTCC 60  
Db |||||  
Qy 42 ATGGCCACCTTGTCTCCGAGCAAGCTGTCCAACTGGCCAGCTCCGCTGTCCAAAGTCC 101  
Db |||||  
Qy 61 CAGGCCAAGATAGGGGATGTTCCGAGGATGGGTTTCAAGCGGCCACGGATGAGGAG 120  
Db |||||  
Qy 102 CAGGCCAAGATAGGGGATGTTTCCAGGATGGGGTTTCAAGCGGCCACGGATGAGGAA 161  
Db |||||  
Qy 121 CGGTTGGGCTTCCGCGATTGCGAGCCTCCACTTTGAGCACCGCCAGGGGCTCGAGATG 180  
Db |||||  
Qy 162 CGGTTGGGCTTCCGCGATTGCGAGCCTCCACTTTGAGCACCGCCAGGGGCTCGAGATG 221  
Db |||||  
Qy 181 GACATCTCTGAAAGCGAGGAGAGCCCTTCCGCGGACGAGGCGCTGAAGCGCCGCTCGAG 240  
Db |||||  
Qy 222 GACATCTCTGAAAGCGAGGAGAGCCCTTCCGCGGACGAGGCGCTGAAGCGCCGCTCGAG 281  
Db |||||  
Qy 241 GGAGACATTCATATTCAGCGAGGAGCGGAGCTCTTTCGCGCCCTCCGGCTCCAAAGAC 300  
Db |||||  
Qy 282 GGAGACATTCATATTCAGCGCGGGGGCTCTCTGCGCCCTCTGGCTTCCAAAGACCGAG 341  
Db |||||  
Qy 301 CAGTTGGGAGTGTGGCGAATTCGGGGGCCACAGAGCCCAAAATCAAGCGCTGGGAG 360  
Db |||||  
Qy 342 GCCGTGGAGCTGTGGGGAGTTGGGGGGTTCACACAAACCAAGATCAAGCGGCTGGGAA 401  
Db |||||  
Qy 361 GCAGGCTGGAACTGACCAAGCCATCCAGGCGATGTTGCTGCTGGGCTTACCTTACGCC 420  
Db |||||  
Qy 402 CGGGCTGGAACTGACCAATGCCATTCAGGCGATGTTGCTGCTGGGCTTACCTTACGCC 461  
Db |||||  
Qy 421 ATCTGCAAGCGGCTACCTGGGTTGTTCTCATATCTTTCGCGCGCTTGTGTGCTGC 480  
Db |||||  
Qy 462 ATCTTCCACGCGGCTACCTGGGTTGTTCTCATCATCTTTCGCGCGAGTGTGTGCTGC 521  
Db |||||  
Qy 481 TACACCGGAGATCTCATCGCTGCTGTACAGGAGGATGAAGACGGGAGGTGGTG 540  
Db |||||  
Qy 522 TACACCGGAGATCTCATCGCTGCTGTACAGGAGGATGAAGACGGGAGGTGGTG 581  
Db |||||  
Qy 541 CGCGTGGGGAGCTGTACGTTGGCCATAGCAACCGCTGCTGCGCCCGCGCTTCCCAACG 600  
Db |||||  
Qy 582 CGCGTGGGGAGCTGTATGTGGCCATAGCTAAGCCATGCTGCGCTCTCGATTTCCCAAC 641  
Db |||||  
Qy 601 CTGGGCGGCGAGTGTGTAAGAGTGAAGCAGATCATCGAGCTGTGTGATGAGCGTGCATCT 660  
Db |||||  
Qy 642 CTGGGCGGCGGCTGCTCAATGTGGCGCAGATCATCGAGCTGTGTGATGAGCGTGTATCT 701  
Db |||||  
Qy 661 TACGTGTGTGTGAGTGCACCTCATGTACAGAGCTTCCCGGGCTGCGGTGTGCGAG 720  
Db |||||  
Qy 702 TACGTGTGTGTGAGTGCACCTCATGTACAGAGCTTCCCGGGCTGCGGTGTGCGAG 761  
Db |||||  
Qy 721 AAGTCTGTGTCCATTTATTCGCGCAGCGGCTGTGCTGCTGCTGCTTCCCTTAAAGACTC 780  
Db |||||  
Qy 762 AAGTCTGTGTCCATATAGCAGCGGCTGTGCTGCTGCTGCTGCTTCCCTTAAAGACTC 821  
Db |||||  
Qy 781 AAGCGGCTGTCCAAAGTTCAGTCTGCTGTGCACTCTGCGCCCACTTTCGTCATCAATCTCT 840  
Db |||||  
Qy 822 AAGCGGCTGTCCAAAGTTCAGTCTGCTGTGACGCTGCGCCCACTTTCGTCATCAATCTCT 881  
Db |||||

Qy 841 GTCATAGCTACTGTCTATCGCGGCGCGAGCTGGGCTGGGAGAGGTCAAGTTCTTAC 900  
Db |||||  
Qy 882 GTCATAGCTACTGTCTATCGCGGCGCGAGTATGGGCTGGGAGAGGTCAAGTTCTTAC 941  
Db |||||  
Qy 901 ATCGACGCTCAAGAGTTCCTCCATCTCCATTTGGCATCATCGTGTTCAGTACACGCTTCTCAG 960  
Db |||||  
Qy 942 ATCGACGCTCAAGAGTTCCTCCATCTCCATTTGGCATCATCGTGTTCAGTACACGCTTCTCAG 1001  
Db |||||  
Qy 961 ATCTTCTGCTTCTGCTGGAGGCAATATGAGCAGAGCCAGCGAGTTCACATGATGATG 1020  
Db |||||  
Qy 1002 ATCTTCTGCTTCTGCTGGAGGCAATATGAGCAGAGCCAGCGAATTCACATGATGATG 1061  
Db |||||  
Qy 1021 AACTGGACGACATCGAGCTGCTCAAGGGGCTCTTTCGCGCTGCTGCGCTACTCTC 1080  
Db |||||  
Qy 1062 AACTGGACGACATCGAGCTGCTCAAGGGGCTCTTTCGCGCTGCTGCGCTACTCTC 1121  
Db |||||  
Qy 1081 ACCTGGGCGGACGAGACCAAGAGGCTCATCACGGATTAACCTTCCCGGCTTCCATCCGGGCC 1140  
Db |||||  
Qy 1122 ACCTGGGCGGACGAGACCAAGAGGCTCATCACGGATTAACCTTCCCGGCTTCCATCCGGGCC 1181  
Db |||||  
Qy 1141 GTGCTCAACATCTTCTGCTGGGCAAGGCGTGTTCCTATCTCTGCGCATTTCTTCCCGGCC 1200  
Db |||||  
Qy 1182 GTGCTCAACCTTCTGCTGGGCAAGGCGTGTTCCTATCTGCTTTCGCTTCTTCTGCGG 1241  
Db |||||  
Qy 1201 GCTGTCAGGCTGTGGAGAGTCTCTTCCAGGAGGCGAGCGCGCTTTCCTCCCGGCC 1260  
Db |||||  
Qy 1242 GCGCTCAAGTGTGGAGAGTCTCTTCCAGGAGGCGAGTTCGCGCTTCTTCCCGGCC 1301  
Db |||||  
Qy 1261 TGCTTACAGCGGCGAGCGGCGCTTCAAGTCTTGGGGGCTGACGCTGCGCTGCGCGCTCGTC 1320  
Db |||||  
Qy 1302 TGCTTACAGCGGCGAGCGGCTTCAAGTCTTGGGGGCTGACGCTGCGCTGCGCGCTCGTC 1361  
Db |||||  
Qy 1321 GTCTTACGCTGTCTATGCGCATTTATGTCGCGCATTCGCGCTGCTCATGCGGCTCAAC 1380  
Db |||||  
Qy 1362 GTCTTACGCTGTCTATGCGCATTTATGTCGCGCATTCGCGCTGCTCATGCGGCTCAAC 1421  
Db |||||  
Qy 1381 GCGAGCTTACGCGGCGCGGCTTCTTCTGCTGCGGAGCGCTTTCACCTGCGGCTG 1440  
Db |||||  
Qy 1422 GCGAGCTTACGCGGCGCGGCTTCTTCTGCTGCGGAGCGCTTTCACCTGCGGCTG 1481  
Db |||||  
Qy 1441 CTCTGCGGCAAGTGTCTGTCGCGCATTCCTTTCGAGCTGCGCATCTTTCGTCATCGGC 1500  
Db |||||  
Qy 1482 CTCTGCGGCAAGTGTCTGTCGCGCATTCCTTTCGAGCTGCGCATCTTTCGTCATCGGC 1541  
Db |||||  
Qy 1501 GGCATCTGAGGCTGTCGCGCTTCTGTCATCTCCCTGAGGCGCTTTCATGAGGCTTACCGA 1560  
Db |||||  
Qy 1542 GGCATCTGAGGCTGTCGCGCTTCTGTCATCTCCCTGAGGCTTTCATGAGGCTTACCGA 1601  
Db |||||  
Qy 1561 ACCAAGCGGAGGAGTAG 1578  
Db |||||  
Qy 1602 ACCAAGCGGAGGAGTAG 1619  
Db |||||

RESULT 12

AAS06087/c

ID AAS06087 standard; DNA; 1723 BP.

XX AAS06087;

XX AC

XX AC

XX 12-SEP-2001 (first entry)

XX

DE Angiotensin converting enzyme (ACEV) splice variant DNA #87.

XX  
KW Angiotensin converting enzyme splice variant; ACEV; interleukin 6;  
KW granulocyte colony stimulating factor receptor; glucagon; hypertrophy;  
KW platelet-derived endothelial cell growth factor; cardiovascular disease;  
KW cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C; ds;  
KW vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;  
KW myocardial infarction; coronary arterial thrombosis; renal disease;  
KW diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;  
KW multiple sclerosis; immune complex nephritis; deep vein thrombosis;  
KW nonarctoidotic pulmonary granulomatous disease; endothelial abnormality;  
KW vascular disorder; asbestosis.

QY 241 GGAGACATCCATTATCAGCGAGGAGCGAGCTCCTCTCGCGCCCTCCGGCTCCAAGGAC 300  
DB 1383 GGAGACATCCATTATCAGCGAGGAGCGAGCTCCTCTCGCGCCCTCCGGCTCCAAGGAC 1324  
QY 301 CAGGTGGAGGTGGTGGGAAATTCGGGGGCCACGACAAAGCCCAAAATCACCGCGTGGAG 360  
DB 1323 CAGGTGGAGGTGGTGGGAAATTCGGGGGCCACGACAAAGCCCAAAATCACCGCGTGGAG 1264  
QY 361 CGAGGCTGGAGGTGACCAAGCCATCCAGGCGATGTTCTGGGCTACCCCTACGCC 420  
DB 1263 CGAGGCTGGAGGTGACCAAGCCATCCAGGCGATGTTCTGGGCTACCCCTACGCC 1204  
QY 421 ATCTGTCACGGCGGCTACCTGGGGTGTGTTCTCATCATCTTCGCCCGCGTGTGTGCTGC 480  
DB 1203 ATCTGTCACGGCGGCTACCTGGGGTGTGTTCTCATCATCTTCGCCCGCGTGTGTGCTGC 1144  
QY 481 TACACCGGCAAGATCCTCATCGCTGCTGTGACGAGGAGAAATGAAGACGGGAGGTGTG 540  
DB 1143 TACACCGGCAAGATCCTCATCGCTGCTGTGACGAGGAGAAATGAAGACGGGAGGTGTG 1084  
QY 541 CGGTCGGGACTCTAGTGGCCATAGCCACGCTGCTGCGCCCGCGGCTTCCCAAGC 600  
DB 1083 CGGTCGGGACTCTAGTGGCCATAGCCACGCTGCTGCGCCCGCGGCTTCCCAAGC 1024  
QY 601 CTGGCGCGCGAGTGGTGAACGTAGCGAGATCATCGAGCTGGTGATGACGTGCACTCTG 660  
DB 1023 CTGGCGCGCGAGTGGTGAACGTAGCGAGATCATCGAGCTGGTGATGACGTGCACTCTG 964  
QY 661 TACGTGGTGGTGGAGGCAACCTCATGTACACAGCTTCCCGGGCTGCCCGTGTGCGAG 720  
DB 963 TACGTGGTGGTGGAGGCAACCTCATGTACACAGCTTCCCGGGCTGCCCGTGTGCGAG 904  
QY 721 AAGTCTGTGTCATTATGCGACGCGCTGCTGCTGCGCTTCCCTTAAGAACCTC 780  
DB 903 AAGTCTGTGTCATTATGCGACGCGCTGCTGCTGCGCTTCCCTTAAGAACCTC 844  
QY 781 AAGGCGGTGTCAGTTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840  
DB 843 AAGGCGGTGTCAGTTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 784  
QY 841 GTCATAGCTACTGTCTATGCGGGCGCGAGCTGGGCTGGGAGAAAGTCAAGTTCATC 900  
DB 783 GTCATAGCTACTGTCTATGCGGGCGCGAGCTGGGCTGGGAGAAAGTCAAGTTCATC 724  
QY 901 ATCGAGCTCAAGAGTTCCTCCATCTCCATGCGATCATGTTGCTGCTGCTGCTGCTGCTG 960  
DB 723 ATCGAGCTCAAGAGTTCCTCCATCTCCATGCGATCATGTTGCTGCTGCTGCTGCTGCTG 664  
QY 961 ATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020  
DB 663 ATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 604  
QY 1021 AACTGGACGACATCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
DB 603 AACTGGACGACATCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 544  
QY 1081 ACTTGGGCGGACGAGACCAAGAGGTGATCAAGGATTAACCTGCGGCTGCTGCTGCTGCTGCTG 1140  
DB 543 ACTTGGGCGGACGAGACCAAGAGGTGATCAAGGATTAACCTGCGGCTGCTGCTGCTGCTGCTG 484  
QY 1141 GTGGTCAACATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200  
DB 483 GTGGTCAACATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 424  
QY 1201 GCTGTGAGGTGCTGAGAGTCTGCTTTCAGGAGGCGAGCGCGCTTTCCTGCGGCTGCTGCGGCT 1260  
DB 423 GCTGTGAGGTGCTGAGAGTCTGCTTTCAGGAGGCGAGCGCGCTTTCCTGCGGCTGCTGCGGCT 364  
QY 1261 TGTACAGGCGGACGCGGCGCTGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320  
DB 363 TGTACAGGCGGACGCGGCGCTGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 304

## RESULT 11

AAS06061

ID AAS06061 standard; DNA; 1648 BP.

XX AAS06061;

AC AAS06061;

DT 12-SEP-2001 (first entry)

XX 12-SEP-2001 (first entry)

DE Angiotensin converting enzyme (ACEV) splice variant DNA #61.

XX Angiotensin converting enzyme splice variant; ACEV; interleukin 6;

KW granulocyte colony stimulating factor receptor; glucagon; hypertrophy;

KW platelet-derived endothelial cell growth factor; cardiovascular disease;

KW cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C; ds;

KW vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;

KW myocardial infarction; coronary arterial thrombosis; renal disease;

KW diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;

KW multiple sclerosis; immune complex nephritis; deep vein thrombosis;

KW nonaroidotic pulmonary granulomatous disease; endothelial abnormality;

KW vascular disorder; asbestosis.

XX Synthetic.

OS WO200136632-A2.

XX 25-MAY-2001.

PN 17-NOV-2000; 2000WO-IL000766.

PD 17-NOV-1999; 99IL-00132978.

XX 10-DEC-1999; 99IL-00133455.

XX (COMP-) COMPUEN LTD.

XX Levine Z, David A, Azar I, Khosravi R, Bernstein J;

XX WPI; 2001-336004/35.

DR P-PSDB; AAU02961.

XX Novel alternative splicing variants e.g. variant of angiotensin

PT converting enzyme (ACEV), useful in identifying candidate compounds

PT capable of binding to the variant and to detect anti-variant antibodies.

XX Claim 1; Page 359; 519pp; English.

XX The sequence represents a DNA encoding an angiotensin converting enzyme

CC splice variant (ACEV) polypeptide. The polypeptides of the invention

CC include variants of granulocyte colony stimulating factor receptor,

CC glucagon, interleukin 6, platelet-derived endothelial cell growth factor,

CC cyclin-dependent kinase inhibitor 1C, cellular tumour antigen P53, and

CC vasoactive intestinal polypeptide receptor 2. The polypeptides and their

CC associated nucleic acids are useful for identification of variant

```
QY 961 ATCTTCTGCTTTCGCTGGAGGCAATATGACAGAGCCAGAGTTCCTCACTGCATGATG 1020
Db 663 ATCTTCTGCTTTCGCTGGAGGCAATATGACAGAGCCAGAGTTCCTCACTGCATGATG 604
QY 1021 AACTGGACGACATCGAGCGCTGCGTCTCAAGGGCCCTCTTGGCGCTCGTGGCTACCTC 1080
Db 603 AACTGGACGACATCGAGCGCTGCGTCTCAAGGGCCCTCTTGGCGCTCGTGGCTACCTC 544
QY 1081 ACCTGGSCCGACGAGACCAAGGAGTCTATCAACGATCACTGCGCGCTCCATCCGGGCC 1140
Db 543 ACCTGGSCCGACGAGACCAAGGAGTCTATCAACGATCACTGCGCGCTCCATCCGGGCC 484
QY 1141 GTGCTCAACATCTTTCTGTGGCCCAAGGCGCTGTGTCTATCTCTGCGCATCTTTGGCC 1200
Db 483 GTGCTCAACATCTTTCTGTGGCCCAAGGCGCTGTGTCTATCTCTGCGCATCTTTGGCC 424
QY 1201 GCTGTCAGGTGCTGGAGAGTGCCTCTTCCAGGAAGGCGCGCTTTTCCCGGCC 1260
Db 423 GCTGTCAGGTGCTGGAGAGTGCCTCTTCCAGGAAGGCGCGCTTTTCCCGGCC 364
QY 1261 TGCTACAGCGCGACGGCGCTGAAGTCTCTGGGGCTGACGTGCGCTCGCGCTCGTC 1320
Db 363 TGCTACAGCGCGACGGCGCTGAAGTCTCTGGGGCTGACGTGCGCTCGCGCTCGTC 304
QY 1321 GTCTTACGCTGCTCATGGCCATTATGTGCGCACTTTCGCGTGTCTATGGGCTCACC 1380
Db 303 GTCTTACGCTGCTCATGGCCATTATGTGCGCACTTTCGCGTGTCTATGGGCTCACC 244
QY 1381 GGCAGCTCACGGCGCGCGCTCTGTCTTCTGCTGCCAGCTCTTTACCTTGGCGCTG 1440
Db 243 GGCAGCTCACGGCGCGCGCTCTGTCTTCTGCTGCCAGCTCTTTACCTTGGCGCTG 184
QY 1441 CTCTGGCGCAAGCTGCTGTGGCAAGTCTTCTTGACGTGCGCATCTTCGTATCGGC 1500
Db 183 CTCTGGCGCAAGCTGCTGTGGCAAGTCTTCTTGACGTGCGCATCTTCGTATCGGC 124
QY 1501 GGCATCTGACGCTGCTCGCGCTTGTGCACTCTCTCGAGGCGCTCATCGAAGCTTACCGA 1560
Db 123 GGCATCTGACGCTGCTCGCGCTTGTGCACTCTCTCGAGGCGCTCATCGAAGCTTACCGA 64
QY 1561 ACCAAGCGGAGGACTAG 1578
Db 63 ACCAAGCGGAGGACTAG 46
```

## RESULT 10

ABK92054/c

ID ABK92054 standard; DNA; 1734 BP.

XX AC ABK92054;

XX XX

XX 14-AUG-2002 (first entry)

XX DE

XX DNA encoding novel GABA transporter-like receptor protein #6.

XX Cancer; metabolic disorder; skin disorder; infectious disease; anorexia;  
XX behavioral disorder; valve disease; endocrine disorder; heart disorder;  
XX blood disorder; anxiety disorder; brain disease; inflammatory disorder;  
XX neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
XX immune disorder; haematopoietic disorder; dyslipidaemia;  
XX metabolic disturbance; obesity; metabolic syndrome X; wasting disorder;  
XX gene therapy; transgenic animal; human; gene; ds.

XX OS Homo sapiens.

XX XX

XX PN WO200229058-A2.

XX XX

XX PD 11-APR-2002.

XX XX

XX PF 05-OCT-2001; 2001WO-US031248.

XX XX

XX PR 05-OCT-2000; 2000US-0238323P.

XX PR

XX 05-OCT-2000; 2000US-0238325P.

PR 06-OCT-2000; 2000US-0238372P.  
PR 06-OCT-2000; 2000US-0238373P.  
PR 06-OCT-2000; 2000US-0238379P.  
PR 06-OCT-2000; 2000US-0238382P.  
PR 06-OCT-2000; 2000US-0238383P.  
PR 06-OCT-2000; 2000US-0238384P.  
PR 06-OCT-2000; 2000US-0238397P.  
PR 06-OCT-2000; 2000US-0238400P.  
PR 06-OCT-2000; 2000US-0238401P.  
PR 06-OCT-2000; 2000US-0238402P.  
PR 14-MAR-2001; 2001US-0275892P.  
PR 08-JUN-2001; 2001US-0296860P.  
XX (CURA-) CURAGEN CORP.  
XX PA  
XX PI Shimkets RA, Taupier RJ, Burgess CE, Zerhusen BD, Mezes PS;  
PI Rastelli L, Malyankar UM, Grosse WM, Alsobrook JP, Lepley DM;  
PI Spytek KA, Li L, Edinger S, Gerlach V, Ellerman K, Macdougall J;  
PI Gunther E, Millet I, Stone D, Smithson G, Szekeres ES;  
XX  
DR WPI; 2002-444103/47.  
XX P-PSDB; ABG61787.  
XX  
PT Novel isolated polypeptide, designated NOVX, useful for treating or  
PT preventing cancer, diabetes, obesity, dyslipidemia, anorexia, and  
PT metabolic, neurodegenerative, immune and hematopoietic disorders.  
XX  
PS Claim 9; Page 64; 316pp; English.  
XX  
XX The invention describes an isolated polypeptide (I), useful in the  
XX manufacture of a medicament for treating a syndrome associated with a  
XX human disease. (I), the polynucleotide encoding it (II) and an antibody  
XX (III) to (I) are useful for treating or preventing cancer, metabolic  
XX disorders, skin disorders, infectious disease, anorexia, behavioral  
XX disorders, valve diseases, endocrine disorders, heart and blood  
XX disorders, anxiety disorders, brain disorders, inflammatory disorders,  
XX neurodegenerative disorder, Alzheimer's disease, Parkinson's disorder,  
XX immune disorders, haematopoietic disorders, and the various  
XX dyslipidaemias, metabolic disturbances associated with obesity, and the  
XX metabolic syndrome X and wasting disorders associated with chronic  
XX diseases. (I), (II) or (III) are useful in screening assays, detection  
XX assays (e.g., chromosomal mapping, tissue typing, forensic biology),  
XX predictive medicine (e.g., diagnostic assays and prognostic assays), and  
XX in methods of treatment (e.g., therapeutic and prophylactic). (I) is  
XX useful as immunogen to produce antibodies immunospecific for (I), to  
XX screen for potential agonist and antagonist compounds, and as bait  
XX protein in a two-hybrid or three-hybrid assay. (II) is useful in gene  
XX therapy, to express (I), to detect (I) mRNA or a genetic lesion in a (I)  
XX gene, and to modulate activity of (I). A cell containing a vector  
XX expressing (I) is useful for producing non-human transgenic animals. This  
XX sequence encodes a novel human polypeptide described in the invention  
XX  
SQ Sequence 1734 BP; 357 A; 514 C; 580 G; 283 T; 0 U; 0 Other;  
Query Match 99.8%; Score 1574.8; DB 6; Length 1734;  
Best Local Similarity 99.9%; Pred. No. 6.9e-311; Mismatches 0; Gaps 0;  
Matches 1576; Conservative 0; Indels 2;

QY 1 ATGGCCACCTTGTCTCGCAGCAGCTGTCCAAAGTGGCCACCGTCCGTGTCCAAACAGTCC 60  
Db 1623 ATGGCCACCTTGTCTCGCAGCAGCTGTCCAAAGTGGCCACCGTCCGTGTCCAAACAGTCC 1564

QY 61 CAGGCCAAGATGAGCGGCATGTTGCGCCAGGATGGGTTTTCAGCGCGCCACGGATAGGAG 120  
Db 1563 CAGGCCAAGATGAGCGGCATGTTGCGCCAGGATGGGTTTTCAGCGCGCCACGGATAGGAG 1504

QY 121 GCGGTGGCTTCGCGCATTTGCGACACCTCGACTTTTGGACACCGCCAGGCGCTTCAGATG 180  
Db 1503 GCGGTGGCTTCGCGCATTTGCGACACCTCGACTTTTGGACACCGCCAGGCGCTTCAGATG 1444

QY 181 GACATCTGAAAGCCGAGGGAGAGCCCTCGCGGGACGAGGGCGCTGAAGCGCCGCTCGAG 240  
Db 1443 GACATCTGAAAGCCGAGGGAGAGCCCTCGCGGGACGAGGGCGCTGAAGCGCCGCTCGAG 1384



AC ABK92053;  
XX 14-AUG-2002 (first entry)  
DE DNA encoding novel GABA transporter-like receptor protein #5.  
XX  
XX Cancer; metabolic disorder; skin disorder; infectious disease; anorexia;  
KW behavioral disorder; valve disease; endocrine disorder; heart disorder;  
KW blood disorder; anxiety disorder; brain disorder; inflammatory disorder;  
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
KW immune disorder; haematopoietic disorder; dyslipidaemia;  
KW metabolic disturbance; obesity; metabolic syndrome X; wasting disorder;  
KW gene therapy; transgenic animal; human; gene; ds.  
XX  
OS Homo sapiens.  
XX  
XX WO200229058-A2.  
XX  
XX 11-APR-2002.  
XX  
XX 05-OCT-2001; 2001WO-US031248.  
XX  
XX 05-OCT-2000; 2000US-0238323P.  
XX 05-OCT-2000; 2000US-0238325P.  
XX 06-OCT-2000; 2000US-0238372P.  
XX 06-OCT-2000; 2000US-0238373P.  
XX 06-OCT-2000; 2000US-0238379P.  
XX 06-OCT-2000; 2000US-0238382P.  
XX 06-OCT-2000; 2000US-0238383P.  
XX 06-OCT-2000; 2000US-0238384P.  
XX 06-OCT-2000; 2000US-0238397P.  
XX 06-OCT-2000; 2000US-0238400P.  
XX 06-OCT-2000; 2000US-0238401P.  
XX 06-OCT-2000; 2000US-0238402P.  
XX 14-MAR-2001; 2001US-0275892P.  
XX 08-JUN-2001; 2001US-0296860P.  
XX  
XX (CURA-) CURAGEN CORP.  
XX  
XX Shinketsu RA, Taupier RJ, Burgess CE, Zerhusen BD, Mezes PS;  
PI Rastelli L, Malyankar UM, Grosse WM, Alsobrook JP, Lepley DM;  
PI Spytek KA, Li L, Edinger S, Gerlach V, Ellerman K, Macdougall J;  
PI Gunther E, Millet I, Stone D, Smithson G, Szekeres ES;  
XX  
XX WPI: 2002-444103/47.  
XX P-PSDB; ABG61786.  
XX  
XX Novel isolated polypeptide, designated NOVX, useful for treating or  
XX preventing cancer, diabetes, obesity, dyslipidaemia, anorexia, and  
XX metabolic, neurodegenerative, immune and hematopoietic disorders.  
XX  
XX Claim 9; Page 63-64; 316pp; English.  
XX  
XX The invention describes an isolated polypeptide (I), useful in the  
XX manufacture of a medicament for treating a syndrome associated with a  
XX human disease. (I), the polynucleotide encoding it (II) and an antibody  
XX (III) to (I) are useful for treating or preventing cancer, metabolic  
XX disorders, skin disorders, infectious disease, anorexia, behavioral  
XX disorders, valve diseases, endocrine disorders, heart and blood  
XX disorders, anxiety disorders, brain disorders, inflammatory disorders,  
XX neurodegenerative disorders, Alzheimer's disease, Parkinson's disease,  
XX immune disorders, haematopoietic disorders, and the various  
XX dyslipidaemias, metabolic disturbances associated with obesity, and the  
XX metabolic syndrome X and wasting disorders associated with chronic  
XX diseases. (I), (II) or (III) are useful in screening assays, detection  
XX assays (e.g., chromosomal mapping, tissue typing, forensic biology),  
XX predictive medicine (e.g., diagnostic assays and prognostic assays), and  
XX in methods of treatment (e.g., therapeutic and prophylactic). (I) is  
XX useful as immunogen to produce antibodies immunospecific for (I), to  
XX screen for potential agonist and antagonist compounds, and as bait  
XX protein in a two-hybrid or three-hybrid assay. (II) is useful in gene  
XX therapy, to express (I), to detect (I) mRNA or a genetic lesion in a (I)  
XX gene, and to modulate activity of (I). A cell containing a vector

CC expressing (I) is useful for producing non-human transgenic animals. This  
CC sequence encodes a novel human polypeptide described in the invention  
XX  
SQ Sequence 1734 BP; 358 A; 515 C; 579 G; 282 T; 0 U; 0 Other;  
Query Match 99.8%; Score 1574.8; DB 6; Length 1734;  
Best Local Similarity 99.9%; Pred. No. 6.9e-311;  
Matches 1576; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Oy 1 ATGCCACCTTGTCTCCGCGAGCAAGCTGTCACAGTGGCCAGCTCGTGTCAACAAAGTCC 60  
Db 1623 ATGCCACCTTGTCTCCGCGAGCAAGCTGTCACAGTGGCCAGCTCGTGTCAACAAAGTCC 1564  
Oy 61 CAGGCCAAGATCAGCGGCAATGTTGCCAGAGTGGGTTTTTCAGGGCGCCACGGATGAGGAG 120  
Db 1563 CAGGCCAAGATCAGCGGCAATGTTGCCAGAGTGGGTTTTTCAGGGCGCCACGGATGAGGAG 1504  
Oy 121 GCGTGGGCTTCGCGCATTTGCGACGACCTCTGACATTTAGACACCGCCAGGGCTTCGAGATG 180  
Db 1503 GCGTGGGCTTCGCGCATTTGCGACGACCTCTGACATTTAGACACCGCCAGGGCTTCGAGATG 1444  
Oy 181 GACATCTGAAAGCCGAGGAGAGCCCTGCGGGGACAGGGCGCTGAAAGCGCCCTCGAG 240  
Db 1443 GACATCTGAAAGCCGAGGAGAGCCCTGCGGGGACAGGGCGCTGAAAGCGCCCTCGAG 1384  
Oy 241 GGAGACATCCATTATCAGCGAGGAGCGGAGCTCTCTCTGCGGCCCTCCGGCTTCCAAGAC 300  
Db 1383 GGAGACATCCATTATCAGCGAGGAGCGGAGCTCTCTCTGCGGCCCTCCGGCTTCCAAGAC 1324  
Oy 301 CAGTGGGAGGTGGTGGGCAATTTGGGGGCGCACAGCAAGCCCAAAATCAGGGCTGGGAG 360  
Db 1323 CAGTGGGAGGTGGTGGGCAATTTGGGGGCGCACAGCAAGCCCAAAATCAGGGCTGGGAG 1264  
Oy 361 CAGGCTGGAACGTGACCAACCGCATCCAGGGCGATGTTCTGCTGGGGCTTACCTTACGCC 420  
Db 1263 CAGGCTGGAACGTGACCAACCGCATCCAGGGCGATGTTCTGCTGGGGCTTACCTTACGCC 1204  
Oy 421 ATCTTGCAACGGCGCTACCTGGGGTGTGTTCTCATCATCTTTCGCGCCGCTGTGTGCTGC 480  
Db 1203 ATCTTGCAACGGCGCTACCTGGGGTGTGTTCTCATCATCTTTCGCGCCGCTGTGTGCTGC 1144  
Oy 481 TACACCGGCAAGATCTCTCATCGTGCCTGTACGAGGAGATGAAGACGGCGAGTGTG 540  
Db 1143 TACACCGGCAAGATCTCTCATCGTGCCTGTACGAGGAGATGAAGACGGCGAGTGTG 1084  
Oy 541 CCGTGGCGGAGCTGTAGTGCATAGCAACGGCTCTGCGGCCCGCGCTTCCCAACG 600  
Db 1083 CCGTGGCGGAGCTGTAGTGCATAGCAACGGCTCTGCGGCCCGCGCTTCCCAACG 1024  
Oy 601 CTGGCGGCGGAGTGGTGAACGTAGCGCAGATCATCGAGCTGGTGATGACGTGCATCTCTG 660  
Db 1023 CTGGCGGCGGAGTGGTGAACGTAGCGCAGATCATCGAGCTGGTGATGACGTGCATCTCTG 964  
Oy 661 TACGTGGTGGTGAAGTGGCAACCTCATGTACACAGCTTCCCGGGGCTGCCGTGTCGAG 720  
Db 963 TACGTGGTGGTGAAGTGGCAACCTCATGTACACAGCTTCCCGGGGCTGCCGTGTCGAG 904  
Oy 721 AAGTCTGTGTCATTATCGCCACGGCGCTGTGCTGCTTGGCGCTTCTCTAAGAACCTC 780  
Db 903 AAGTCTGTGTCATTATCGCCACGGCGCTGTGCTGCTTGGCGCTTCTCTAAGAACCTC 844  
Oy 781 AAGCGCGTGTCCAAAGTTTCAAGTGTGTCGCTTGGGCCCACTTCGTCAATCAATATCTCTG 840  
Db 843 AAGCGCGTGTCCAAAGTTTCAAGTGTGTCGCTTGGGCCCACTTCGTCAATCAATATCTCTG 784  
Oy 841 GTCATAGCTTACTGTCTATCGGGCGCGGAGCTGGGCTTGGGAGAAAGTCAAGTCTTAC 900  
Db 783 GTCATAGCTTACTGTCTATCGGGCGCGGAGCTGGGCTTGGGAGAAAGTCAAGTCTTAC 724  
Oy 901 ATCAGCTCAAGAAAGTTCCCATCTCCATTGGGCATCATCTGTTTTCAGCTTACAGCTCTCAG 960  
Db 723 ATCAGCTCAAGAAAGTTCCCATCTCCATTGGGCATCATCTGTTTTCAGCTTACAGCTCTCAG 664

PT metabolic, neurodegenerative, immune and hematopoietic disorders.  
XX Claim 9; Page 63; 316pp; English.  
XX The invention describes an isolated polypeptide (I), useful in the  
CC manufacture of a medicament for treating a syndrome associated with a  
CC human disease. (I), the polynucleotide encoding it (II) and an antibody  
CC (III) to (I) are useful for treating or preventing cancer, metabolic  
CC disorders, skin disorders, infectious diseases, anorexia, behavioral  
CC disorders, valve diseases, endocrine disorders, heart and blood  
CC neurodegenerative disorder, Alzheimer's disease, Parkinson's disorder,  
CC immune disorders, haematopoietic disorders, and the various  
CC dyslipidaemias, metabolic disturbances associated with obesity, and the  
CC metabolic syndrome X and wasting disorders associated with chronic  
CC diseases. (I), (II) or (III) are useful in screening assays, detection  
CC assays (e.g., chromosomal mapping, tissue typing, forensic biology),  
CC predictive medicine (e.g., diagnostic assays and prognostic assays), and  
CC in methods of treatment (e.g., therapeutic and prophylactic). (I) is  
CC useful as immunogen to produce antibodies immunospecific for (I), to  
CC screen for potential agonist and antagonist compounds, and as bait  
CC protein in a two-hybrid or three-hybrid assay. (II) is useful in gene  
CC therapy, to express (I), to detect (I) mRNA or a genetic lesion in a (I)  
CC gene, and to modulate activity of (I). A cell containing a vector  
CC expressing (I) is useful for producing non-human transgenic animals. This  
CC sequence encodes a novel human polypeptide described in the invention  
XX  
SQ Sequence 1734 BP; 356 A; 515 C; 581 G; 282 T; 0 U; 0 Other;

Query Match 99.8%; Score 1574.8; DB 6; Length 1734;  
Best Local Similarity 99.9%; Pred. No. 6.9e-311;  
Matches 1576; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy	1	ATGGCCACCTTGCTCCGAGCAAGCTGTCCAACTGGGCCAGCTCCGTGTCCAAAGTCC	60
Db	1623	ATGGCCACCTTGCTCCGAGCAAGCTGTCCAACTGGGCCAGCTCCGTGTCCAAAGTCC	1564
Qy	61	CAGGCCAAGATGAGCGGATGTTCCGAGGATGGTTTTCAGGCGGCCAGGATGAGGAG	120
Db	1563	CAGGCCAAGATGAGCGGATGTTCCGAGGATGGTTTTCAGGCGGCCAGGATGAGGAG	1504
Qy	121	CGGTGGGCTTTCGGCATTCGACGACCTCGACTTTTGAGCACCCAGGCGCTGCAGATG	180
Db	1503	CGGTGGGCTTTCGGCATTCGACGACCTCGACTTTTGAGCACCCAGGCGCTGCAGATG	1444
Qy	181	GACATCTGAAAGCCGAGGAGACCCCTGCGGGGACGAGGCGCTGAAGCGCCGCTCGAG	240
Db	1443	GACATCTGAAAGCCGAGGAGACCCCTGCGGGGACGAGGCGCTGAAGCGCCGCTCGAG	1384
Qy	241	GGAGACATCCATTATCAGCGAGGACGCGAGCTCTCTCGGCGCTCCGGCTCCAGGAC	300
Db	1383	GGAGACATCCATTATCAGCGAGGACGCGAGCTCTCTCGGCGCTCCGGCTCCAGGAC	1324
Qy	301	CAGGTGGGAGGTGTGGCGAATTCGGGGGCCACGACAAAGCCAAATACAGCGGTGGGAG	360
Db	1323	CAGGTGGGAGGTGTGGCGAATTCGGGGGCCACGACAAAGCCAAATACAGCGGTGGGAG	1264
Qy	361	GCAGGTGGAAAGCTGAACAAAGCCATCCAGGCAATGTTCTGCTGGGCGCTACCCCTACGCC	420
Db	1263	GCAGGTGGAAAGCTGAACAAAGCCATCCAGGCAATGTTCTGCTGGGCGCTACCCCTACGCC	1204
Qy	421	ATCTCTGACGCGGCTACCTGGGGTGTGTTCTCATCATCTTCGCGCGGTTGTGTCTGC	480
Db	1203	ATCTCTGACGCGGCTACCTGGGGTGTGTTCTCATCATCTTCGCGCGGTTGTGTCTGC	1144
Qy	481	TACACCGGCAAGATCTCATCGGTGCTGTACGAGGAGATGAAGACGCGAGGTGGTG	540
Db	1143	TACACCGGCAAGATCTCATCGGTGCTGTACGAGGAGATGAAGACGCGAGGTGGTG	1084
Qy	541	CGCGTGGGAGCTCGTACGTGGCCATAGCAAAGCTGTGCGGCCCGCGCTTCCCAACG	600
Db	1083	CGCGTGGGAGCTCGTACGTGGCCATAGCAAAGCTGTGCGGCCCGCGCTTCCCAACG	1024

RESULT 9  
ABK92053/C  
ID ABK92053 standard; DNA; 1734 BP.  
XX

QY 241 GGAGACATCATATATCAGCGAGCGCAGCGAGCTCTCTCGCGCCCTCGGCTCCAGGAC 300  
DB 1383 GGAGACATCATATATCAGCGAGCGCAGCGAGCTCTCTCGCGCCCTCGGCTCCAGGAC 1324  
QY 301 CAGGTGGAGGTGGTGGCGAATTCGGGGGCCACGACAGCCCAAAATCAACGGCGTGGAG 360  
DB 1323 CAGGTGGAGGTGGTGGCGAATTCGGGGGCCACGACAGCCCAAAATCAACGGCGTGGAG 1264  
QY 361 CGAGGCTGAAGCGTGAAGCGCATCCAGGCGATGTTGCTGGGCTACCTACGCC 420  
DB 1263 CGAGGCTGAAGCGTGAAGCGCATCCAGGCGATGTTGCTGGGCTACCTACGCC 1204  
QY 421 ATCTGTCAGCGGGGTACCTGGGGTGTGTTCTCATCATCTTCGCGCGCGTGTGCTGC 480  
DB 1203 ATCTGTCAGCGGGGTACCTGGGGTGTGTTCTCATCATCTTCGCGCGCGTGTGCTGC 1144  
QY 481 ATACCGCGAAGATCTCATCTCGGCTGCTGTAACGAGGAATGAAGACGGCGAGTGTG 540  
DB 1143 ATACCGCGAAGATCTCATCTCGGCTGCTGTAACGAGGAATGAAGACGGCGAGTGTG 1084  
QY 541 CGCGTGGGAGCTCGTACGTCGCTAGCAAGCGCTGCTGCGCGCCCGCTTCCCAAG 600  
DB 1083 CGCGTGGGAGCTCGTACGTCGCTAGCAAGCGCTGCTGCGCGCCCGCTTCCCAAG 1024  
QY 601 CTGGCGCGGAGTGTGTAACGTAGCGCAGATCATCGAGCTGCTGATGACGTGCATCTG 660  
DB 1023 CTGGCGCGGAGTGTGTAACGTAGCGCAGATCATCGAGCTGCTGATGACGTGCATCTG 964  
QY 661 TACGTGGTGGTGAAGTGGCAACCTCATGTACACAGCTTCCGGGGCTGCCCTGTGCGAG 720  
DB 963 TACGTGGTGGTGAAGTGGCAACCTCATGTACACAGCTTCCGGGGCTGCCCTGTGCGAG 904  
QY 721 AGTCTGTGTCATATTCACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780  
DB 903 AGTCTGTGTCATATTCACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 844  
QY 781 AAGGCGCTGTCCAAAGTTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840  
DB 843 AAGGCGCTGTCCAAAGTTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 784  
QY 841 GTCATAGCTACTGTCTATCGCGGGCGCGGAGCTGGGCTGGGAGAGGTCAAGTCTTAC 900  
DB 783 GTCATAGCTACTGTCTATCGCGGGCGCGGAGCTGGGCTGGGAGAGGTCAAGTCTTAC 724  
QY 901 ATCGAGCTCAAGAAGTTCCTCCATTCATGTCATCATGTCGTGTCAGCTACAGCTCAG 960  
DB 723 ATCGAGCTCAAGAAGTTCCTCCATTCATGTCATCATGTCGTGTCAGCTACAGCTCAG 664  
QY 961 ATCTTCTGCTTCTGCTGGAGGCAATATGACAGCGCCAGCGAGTTCACCTGCATGATG 1020  
DB 663 ATCTTCTGCTTCTGCTGGAGGCAATATGACAGCGCCAGCGAGTTCACCTGCATGATG 604  
QY 1021 AACTGAGCGACATCGAGCTGCTGCTCAAGGGCTCTTCGCGCTGCTGCGCTACCTC 1080  
DB 603 AACTGAGCGACATCGAGCTGCTGCTCAAGGGCTCTTCGCGCTGCTGCGCTACCTC 544  
QY 1081 ACTTGGGCGAGAGACCGAGAGGTTCATCAGGATAAAGCTGCGGCTCCATCCGCGCC 1140  
DB 543 ACTTGGGCGAGAGACCGAGAGGTTCATCAGGATAAAGCTGCGGCTCCATCCGCGCC 484  
QY 1141 GTGGTCAACATCTTCTGCTGGGCAAGGGGCTGTTGCTCTATCTGCTGCTGCTGCTGCTG 1200  
DB 483 GTGGTCAACATCTTCTGCTGGGCAAGGGGCTGTTGCTCTATCTGCTGCTGCTGCTGCTG 424  
QY 1201 GCTGTGAGGTGCTGAGAGTCTGCTTCCAGGAAGGAGCGCGGCTCTTTTCCGCGCC 1260  
DB 423 GCTGTGAGGTGCTGAGAGTCTGCTTCCAGGAAGGAGCGCGGCTCTTTTCCGCGCC 364  
QY 1261 TGCTACAGGGCGACCGGCGCTGAGAGTCTGCGGGCTGACCGTGGCTGCGGCTGCTGCTG 1320  
DB 363 TGCTACAGGGCGACCGGCGCTGAGAGTCTGCGGGCTGACCGTGGCTGCGGCTGCTGCTG 304  
QY 1321 GTCTTACGCTGCTCATGGCCATTTATGTGCGCGCACTTTCGCGCTGCTCATGGGCTCACC 1380

DB 303 GTCTTACGCTGCTCATGGCCATTTATGTGCGCACTTTCGCGCTCATGGGCTCACC 244  
QY 1381 GGCAGCCTCAAGGGGGCGGCTCTGTTTCTGTCGCCAGGCTCTTTTACCTGCGGCTG 1440  
DB 243 GGCAGCCTCAAGGGGGCGGCTCTGTTTCTGTCGCCAGGCTCTTTTACCTGCGGCTG 184  
QY 1441 CTCTGGGCAAGCTGCTGTGGCACCAAGTCTTCTTCGACGTCGCCATCTTCTGTCATCGGC 1500  
DB 183 CTCTGGGCAAGCTGCTGTGGCACCAAGTCTTCTTCGACGTCGCCATCTTCTGTCATCGGC 124  
QY 1501 GGCATCTGCAAGCTGCTGCGCTTCTGTCATCTTCTGCGAGGCGCTCATCGAAGCCTACCGA 1560  
DB 123 GGCATCTGCAAGCTGCTGCGCTTCTGTCATCTTCTGCGAGGCGCTCATCGAAGCCTACCGA 64  
QY 1561 ACCAAGCGGAGGACTAG 1578  
DB 63 ACCAAGCGGAGGACTAG 46  
RESULT B  
ABK92052/c  
ID ABK92052 standard; DNA; 1734 BP.  
XX AC ABK92052;  
XX XX  
XX DT 14-AUG-2002 (first entry)  
XX XX  
DE DNA encoding novel GABA transporter-like receptor protein #4.  
KW Cancer; metabolic disorder; skin disorder; infectious disease; anorexia;  
KW behavioral disorder; valve disease; endocrine disorder; heart disorder;  
KW blood disorder; anxiety disorder; brain disorder; inflammatory disorder;  
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
KW immune disorder; haematopoietic disorder; dyslipidaemia;  
KW metabolic disturbance; obesity; metabolic syndrome X; wasting disorder;  
KW gene therapy; transgenic animal; human; gene; ds.  
XX OS Homo sapiens.  
XX XX  
XX PN WO200229058-A2.  
XX XX  
XX PD 11-APR-2002.  
XX XX  
XX PF 05-OCT-2001; 2001WO-US031248.  
XX XX  
XX PR 05-OCT-2000; 2000US-0238323P.  
XX PR 05-OCT-2000; 2000US-0238325P.  
XX PR 06-OCT-2000; 2000US-0238372P.  
XX PR 06-OCT-2000; 2000US-0238373P.  
XX PR 06-OCT-2000; 2000US-0238379P.  
XX PR 06-OCT-2000; 2000US-0238382P.  
XX PR 06-OCT-2000; 2000US-0238383P.  
XX PR 06-OCT-2000; 2000US-0238384P.  
XX PR 06-OCT-2000; 2000US-0238397P.  
XX PR 06-OCT-2000; 2000US-0238400P.  
XX PR 06-OCT-2000; 2000US-0238401P.  
XX PR 06-OCT-2000; 2000US-0238402P.  
XX PR 14-MAR-2001; 2001US-0275892P.  
XX PR 08-JUN-2001; 2001US-0296860P.  
XX XX  
XX PA (CURA-) CURAGEN CORP.  
XX XX  
XX PI Shinkets RA, Taupier RJ, Burgess CE, Zerhusen BD, Mezes PS;  
XX PI Rastelli L, Malyankar UM, Grosse WM, Alsbrook JP, Lepley DM;  
XX PI Spytek KA, Li L, Edinger S, Gerlach V, Ellerman K, Macdougall J;  
XX PI Gunther E, Millet I, Stone D, Smithson G, Szekeres ES;  
XX WPI: 2002-444103/47.  
XX DR P-PSDB; ABG61785.  
XX PT Novel isolated polypeptide, designated NOVX, useful for treating or preventing cancer, diabetes, obesity, dyslipidemia, anorexia, and



DT 09-APR-2002 (first entry)  
XX Human vesicular GABA transporter (VGAT) gene sequence.  
DE  
XX Human; vesicular GABA transporter; VGAT; gamma aminobutyric acid; GABA;  
KW nootropic; neuroprotective; antitumor; anticonvulsant; analgesic;  
KW antinflammatory; neuroleptic; chromosome localisation; antidepressant;  
KW antinflammatory; tranquiliser; vaccine; VGAT polypeptide modulator;  
KW screening assay; agonist; antagonist; schizophrenia; epilepsy;  
KW depression; learning disorder; pain; cognitive disorder;  
KW neurodegenerative disease; multiple sclerosis; dementia;  
KW Alzheimer's disease; Parkinson's disease; Crohn's disease;  
KW ulcerative colitis; dyspepsia; irritable bowel syndrome; hyperactivity;  
KW anxiety disorder; sleeping disorder; alcoholism; muscular disorder;  
KW tremor; headache; migraine; immunogen; membrane bound receptor;  
KW soluble receptor; hybridisation probe; primer; diagnostic assay;  
KW tissue expression; transgenic animal; chromosome 20q12-20q13; ds.  
XX  
OS Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FH 100..1677  
FT CDS /\*tag= a  
FT /product= "Human VGAT"  
FT  
XX WO200173015-A1.  
XX  
XX 04-OCT-2001.  
XX  
XX 23-MAR-2001; 2001WO-BP003350.  
XX  
XX 27-MAR-2000; 2000EP-00106581..  
XX (MERE ) MERCK PATENT GMBH.  
XX  
XX Rippmann F, Duecker K;  
PI WPI: 2001-616495/71.  
XX P-PSDB; AAU09936.  
XX  
XX New polypeptide for diagnosing, treating neurodegenerative diseases,  
PT epilepsy, pain, sleeping, anxiety disorder and identifying modulators  
PT comprises the human vesicular gamma aminobutyric acid transporter  
PT polypeptide.  
XX  
XX Claim 5; Page 31-33; 39pp; English.  
XX  
XX The invention relates to a new polypeptide comprising a fully defined  
CC sequence of 525 amino acids as given in the specification, or a  
CC polypeptide comprising a sequence 95% homologous to the fully defined  
CC sequence of the human vesicular gamma aminobutyric acid transporter,  
CC (VGAT). The invention also relates to fragments or variants of VGAT,  
CC encoded by a polynucleotide comprising a fully defined sequence of 1800  
CC base pairs as given in the specification. The VGAT polypeptide is useful  
CC in screening assays to identify compounds that stimulate or inhibit the  
CC function or level of the polypeptide. VGAT polypeptides, polynucleotides  
CC and the agonists and antagonists identified by the above method are  
CC useful for treating schizophrenia, epilepsy, depression, learning  
CC disorders, cognitive disorders, neurodegenerative diseases, multiple  
CC sclerosis, dementia, Alzheimer's, Parkinson's disease, Crohn's disease,  
CC ulcerative colitis, dyspepsia, irritable bowel syndrome, hyperactivity,  
CC anxiety disorder, sleeping disorder, alcoholism, muscular disorders e.g.  
CC tremor, pain, headache and migraine. The VGAT polypeptides and  
CC polynucleotides are also useful as vaccines and as immunogens to produce  
CC antibodies which are useful for treating diseases. VGAT polypeptides can  
CC be used to identify membrane bound or soluble receptors and as  
CC hybridisation probes for cDNA and genomic DNA, as primers for nucleic  
CC acid amplification reactions to isolate full-length cDNAs and genomic  
CC clones encoding VGAT polypeptides, in diagnostic assays by detecting  
CC mutations in the associated gene, for chromosome localisation studies,  
CC tissue expression studies and for producing transgenic animals useful in  
CC drug discovery and target validation. The present sequence encodes the  
CC human vesicular gamma aminobutyric acid transporter (VGAT) protein of the

CC invention. This gene maps to human chromosome 20q12-20q13  
XX  
SQ Sequence 1800 BP; 290 A; 613 C; 525 G; 372 T; 0 U; 0 Other;  
Query Match 99.9%; Score 1576.4; DB 5; Length 1800;  
Best Local Similarity 99.9%; Pred. No. 3.3e-311;  
Matches 1577; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 ATGGCCACTTGTCTCCGCGAGCAAGCTGTCACCAAGTGGCCAGCTCGTGTGTCACCAAGTCC 60  
DB 100 ATGGCCACTTGTCTCCGCGAGCAAGCTGTCACCAAGTGGCCAGCTCGTGTGTCACCAAGTCC 159  
QY 61 CAGGCCAAGATGAGCGGCGATGTTCCCAAGATGGTGTTCAGGGGGCCACCGATGAGGAG 120  
DB 160 CAGGCCAAGATGAGCGGCGATGTTCCCAAGATGGTGTTCAGGGGGCCACCGATGAGGAG 219  
QY 121 CGGTGGGCTTCGGCGCATTTGCGACCTGACCTTTGAGCACCGCCAGGGCTTCAGATG 180  
DB 220 CGGTGGGCTTCGGCGCATTTGCGACCTGACCTTTGAGCACCGCCAGGGCTTCAGATG 279  
QY 181 GACATCTGAAAGCCGAGGAGAGCCCTGCGGGAGAGGGCGCTGAAGCGCCCTCGAG 240  
DB 280 GACATCTGAAAGCCGAGGAGAGCCCTGCGGGAGAGGGCGCTGAAGCGCCCTCGAG 339  
QY 241 GGAGACATCCATTATCAGCGAGGAGCGAGCTCTCTGCGGCCCTCCGGTCCAAAGGAC 300  
DB 340 GGAGACATCCATTATCAGCGAGGAGCGAGCTCTCTGCGGCCCTCCGGTCCAAAGGAC 399  
QY 301 CAGGTGGAGGTGGTGGCGAATTGGGGGGCCACGACCAAGCCCAAAATCACGGCGTGGGAG 360  
DB 400 CAGGTGGAGGTGGTGGCGAATTGGGGGGCCACGACCAAGCCCAAAATCACGGCGTGGGAG 459  
QY 361 CGAGCTCGAAAGCTGACCAAGCCCATCCAGGGCATGTTCTGCTGGGCTTACCTACGCC 420  
DB 460 CGAGCTCGAAAGCTGACCAAGCCCATCCAGGGCATGTTCTGCTGGGCTTACCTACGCC 519  
QY 421 ATCTTGACGGCGCTACCTGGGGTTGTTTCTCATCATCTTTCGCCCGCTTGTGTCTGC 480  
DB 520 ATCTTGACGGCGCTACCTGGGGTTGTTTCTCATCATCTTTCGCCCGCTTGTGTCTGC 579  
QY 481 TACACGGCAAGATCCTCATCGCTGCTGTACGAGGAGATGAAGCGGCGGTGGT 540  
DB 580 TACACGGCAAGATCCTCATCGCTGCTGTACGAGGAGATGAAGCGGCGGTGGT 639  
QY 541 CGCTGCGGGACTCGTACGTGCGCCATAGCACAAGCCTGCTGCGCCCGCCGCTTCCCAAG 600  
DB 640 CGCTGCGGGACTCGTACGTGCGCCATAGCACAAGCCTGCTGCGCCCGCCGCTTCCCAAG 699  
QY 601 CTGGCGCGCGAGTGGTGAACGTAGCGAGATCATCGAGCTGGTGAACGTAGCATCTCTG 660  
DB 700 CTGGCGCGCGAGTGGTGAACGTAGCGAGATCATCGAGCTGGTGAACGTAGCATCTCTG 759  
QY 661 TACGTGGTGGTGGGCAACCTCATGTACACAGCTTCCCGGGCTGCGCGTGGCGAG 720  
DB 760 TACGTGGTGGTGGGCAACCTCATGTACACAGCTTCCCGGGCTGCGCGTGGCGAG 819  
QY 721 AAGTCTGTGCTCATTTATCGCCACCGCTGCTGCTGCTGCTGCTGCTTAAAGAACCTC 780  
DB 820 AAGTCTGTGCTCATTTATCGCCACGCGCTGCTGCTGCTGCTGCTTAAAGAACCTC 879  
QY 781 AAGGCGGTGTCCAAGTTCAGTCTGTGTGACATCTGTGGCCCATCTCGTCAATCAATCTCTG 840  
DB 880 AAGGCGGTGTCCAAGTTCAGTCTGTGTGACATCTGTGGCCCATCTCGTCAATCAATCTCTG 939  
QY 841 GTCATAGCTTACTGTCTATCGCGGGCGCGAGCTGGGCTTGGAGAGGTCAAGTCTAC 900  
DB 940 GTCATAGCTTACTGTCTATCGCGGGCGCGAGCTGGGCTTGGAGAGGTCAAGTCTAC 999  
QY 901 ATCGAGCTCAAGAAGTTCCTCCATTCCTCATCATCTGTTTCTGCTGCTGCTGCTGCTGCT 960  
DB 1000 ATCGAGCTCAAGAAGTTCCTCCATTCCTCATCATCTGTTTCTGCTGCTGCTGCTGCTGCT 1059  
QY 961 ATCTTCTCTGCTTTCGCTGGAGGGCAATATGACAGCAGCCAGCGAGTTCACCTGATGATG 1020



Db 1383 GGAGATCCATTATCAGCGAGCGAGAGTCTCTGCGCGCTCCGGCTCCAAGGAC 1324  
Qy 301 CAGGTGGAGGTGGTGGCGAATTCGGGGGCGCACAGCAAGCCCAAAATCAAGCGGTGGGAG 360  
Db 1323 CAGGTGGAGGTGGTGGCGAATTCGGGGGCGCACAGCAAGCCCAAAATCAAGCGGTGGGAG 1264  
Qy 361 CGAGGCTGGAAAGTGAACCAAGCCCATCAGGGCATGTTGCTGCTGGGCTACCCCTACGCC 420  
Db 1263 GCAGGCTGGAAAGTGAACCAAGCCCATCAGGGCATGTTGCTGCTGGGCTACCCCTACGCC 1204  
Qy 421 ATCTGCAAGCGGCTACTCTGGGGTGTGTTCTCATCATCTTCGCGCGCGTGTGTGCTGC 480  
Db 1203 ATCTGCAAGCGGCTACTCTGGGGTGTGTTCTCATCATCTTCGCGCGCGTGTGTGCTGC 1144  
Qy 481 TACACCGCAAGATCCTCATCGCTGCTGTACAGAGGAGATGAAGACGGGAGGTGTG 540  
Db 1143 TACACCGCAAGATCCTCATCGCTGCTGTACAGAGGAGATGAAGACGGGAGGTGTG 1084  
Qy 541 CGCTGCGGAGCTCGTAGTGCCCATAGCCAAAGCGCTGCTGCGCGCCCGCTTCCCAAG 600  
Db 1083 CGCGTGGGACTCGTAGTGCCCATAGCCAAAGCGCTGCTGCGCGCCCGCTTCCCAAG 1024  
Qy 601 CTGGGCGGCGAGTGGTGAACCTAGCGGAGATCATCGAGCTGGTGAATGACGTGATCTCG 660  
Db 1023 CTGGGCGGCGAGTGGTGAACCTAGCGGAGATCATCGAGCTGGTGAATGACGTGATCTCG 964  
Qy 661 TACGTGGTGGTGAAGTGGCAACCTCATGTACACAGCTTCCCGGGCTGCCGTGCGAG 720  
Db 963 TACGTGGTGGTGAAGTGGCAACCTCATGTACACAGCTTCCCGGGCTGCCGTGCGAG 904  
Qy 721 AAGTCTGCTGCTATATGCCACGCGCTGCTGCTGCTGCGCTTCCCTTAAGAACCTC 780  
Db 903 AAGTCTGCTGCTATATGCCACGCGCTGCTGCTGCTGCGCTTCCCTTAAGAACCTC 844  
Qy 781 AAGCGCGTGCAGTTCAGTTCGTGTGCTGCTGCTGCGCTTCCCTTAAGAACCTC 840  
Db 843 AAGCGCGTGCAGTTCAGTTCGTGTGCTGCTGCTGCGCTTCCCTTAAGAACCTC 784  
Qy 841 GTCATAGCTTACTGTCTATCGCGGCGCGCGAGCTGGGCTGGAGAGTCAAGTTCATC 900  
Db 783 GTCATAGCTTACTGTCTATCGCGGCGCGCGAGCTGGGCTGGAGAGTCAAGTTCATC 724  
Qy 901 ATCGAGCTCAAGAGTTCCTCCATCTCATTTGGCATCATGTTTTCAGCTACAGCTCAG 960  
Db 723 ATCGAGCTCAAGAGTTCCTCCATCTCATTTGGCATCATGTTTTCAGCTACAGCTCAG 664  
Qy 961 ATCTTCTGCTGCTGCTGGAGGCAATATGCAAGCGCCAGCGAGTTCACATGATG 1020  
Db 663 ATCTTCTGCTGCTGCTGGAGGCAATATGCAAGCGCCAGCGAGTTCACATGATG 604  
Qy 1021 AACTGGACGACATCGACGCTGCTGCTCAAGGGCTTTCGCGCTGCTGCGCTACCTC 1080  
Db 603 AACTGGACGACATCGACGCTGCTGCTCAAGGGCTTTCGCGCTGCTGCGCTACCTC 544  
Qy 1081 AACTGGGCGGACGAGACCAAGAGGTTCATCAGGATAACTGCGCGGCTTCCATCGCGCC 1140  
Db 543 AACTGGGCGGACGAGACCAAGAGGTTCATCAGGATAACTGCGCGGCTTCCATCGCGCC 484  
Qy 1141 GTGGTCAACATCTTCTGGTGCCCAAGGCGCTGTTGCTATCTCTGCGCATCTTTGCG 1200  
Db 483 GTGGTCAACATCTTCTGGTGCCCAAGGCGCTGTTGCTATCTCTGCGCATCTTTGCG 424  
Qy 1201 GCTGTGAGGTGCTGAGAGTGGCTTTCAGGAGGCGAGCGCGCTTTTTCGCGCC 1260  
Db 423 GCTGTGAGGTGCTGAGAGTGGCTTTCAGGAGGCGAGCGCGCTTTTTCGCGCC 364  
Qy 1261 TGCTACAGCGGACGCGCGCTTGAAGTCTTGGGGCTGACGCTGCGCTGCGCTGCTC 1320  
Db 363 TGCTACAGCGGACGCGCGCTTGAAGTCTTGGGGCTGACGCTGCGCTGCGCTGCTC 304  
Qy 1321 GTCTTACGCTGCTCATGGCCATTTATGTGCGGCACTTCGCGCTGCTCATGGGCTTACC 1380  
Db 303 GTCTTACGCTGCTCATGGCCATTTATGTGCGGCACTTCGCGCTGCTCATGGGCTTACC 244

RESULT 5  
ABK92049

ID ABK92049 standard; DNA; 1763 BP.

XX ABK92049;

AC ABK92049;  
XX  
DT 14-AUG-2002 (first entry)  
XX

XX DNA encoding novel GABA transporter-like receptor protein #1.

XX Cancer; metabolic disorder; skin disorder; infectious disease; anorexia;  
KW behavioral disorder; valve disease; endocrine disorder; heart disorder;  
KW blood disorder; anxiety disorder; brain disorder; inflammatory disorder;  
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
KW immune disorder; hematopoietic disorder; dyslipidaemia;  
KW metabolic disturbance; obesity; metabolic syndrome X; wasting disorder;  
KW gene therapy; transgenic animal; human; gene; ds.

XX Homo sapiens.

XX WO200229058-A2.

XX PD 11-APR-2002.

XX PF 05-OCT-2001; 2001WO-US031248.

XX PR 05-OCT-2000; 2000US-0238323P.

XX PR 05-OCT-2000; 2000US-0238325P.

XX PR 06-OCT-2000; 2000US-0238372P.

XX PR 06-OCT-2000; 2000US-0238373P.

XX PR 06-OCT-2000; 2000US-0238379P.

XX PR 06-OCT-2000; 2000US-0238382P.

XX PR 06-OCT-2000; 2000US-0238383P.

XX PR 06-OCT-2000; 2000US-0238384P.

XX PR 06-OCT-2000; 2000US-0238397P.

XX PR 06-OCT-2000; 2000US-0238400P.

XX PR 06-OCT-2000; 2000US-0238401P.

XX PR 14-MAR-2001; 2001US-0275892P.

XX PR 08-JUN-2001; 2001US-0296860P.

XX (CURA-) CURAGEN CORP.

XX PA Shinketsu RA, Taupier RJ, Burgess CE, Zerhusen BD, Mezes PS;

XX PI Rastelli L, Malyankar UM, Grosse WM, Alsobrook JP, Lepley DM;

XX PI Spytek KA, Li L, Edinger S, Gerlach V, Ellerman K, Macdougall J;

XX PI Gunther E, Millet I, Stone D, Smithson G, Szekeres ES;

XX WPI; 2002-444103/47.

XX P-PSDB; ABG61782.

XX Novel isolated polypeptide, designated NOVX, useful for treating or

XX preventing cancer, diabetes, obesity, dyslipidemia, anorexia, and

XX metabolic, neurodegenerative, immune and hematopoietic disorders.



QY 1021 AACTGGACGACATCGCAGCTCGTCTCAAGGGCCCTCTTGGCGCTCGCTACCTC 1080  
Db 1429 AACTGGACGACATCGCAGCTCGTCTCAAGGGCCCTCTTGGCGCTCGCTACCTC 1488  
QY 1081 ACCTGGCCGACGAGACCAAGGAGGTTCATCACGGATAACCTGCGCGCTCCATCCGCGCC 1140  
Db 1489 ACCTGGCCGACGAGACCAAGGAGGTTCATCACGGATAACCTGCGCGCTCCATCCGCGCC 1548  
QY 1141 GTGGTCAACATCTTTCTGGTGGCCAGGCGCTGTGTCTTATCTCTGCGCATCTTTTGGC 1200  
Db 1549 GTGGTCAACATCTTTCTGGTGGCCAGGCGCTGTGTCTTATCTCTGCGCATCTTTTGGC 1608  
QY 1201 GCTGTCCAGGTCTGGAGAGTCTCTTCCAGGAGGCGAGCGCGCTTTTTCGCGCC 1260  
Db 1609 GCTGTCCAGGTCTGGAGAGTCTCTTCCAGGAGGCGAGCGCGCTTTTTCGCGCC 1668  
QY 1261 TGCTACAGCGGACGCGCGCCCTGAAGTCTCTGGGGCTGACGCTGCGCTCGCGCTCGTC 1320  
Db 1669 TGCTACAGCGGACGCGCGCCCTGAAGTCTCTGGGGCTGACGCTGCGCTCGCGCTCGTC 1728  
QY 1321 GTCTTACGCTGTCTATGGCCATTTATGTGCGGCACTTTCGCGTGTCTATGGGCTTCAAC 1380  
Db 1729 GTCTTACGCTGTCTATGGCCATTTATGTGCGGCACTTTCGCGTGTCTATGGGCTTCAAC 1788  
QY 1381 GGCAGCTCTACGGCGCGCGCCCTCTTTCTTCTGCTGCCAGCCTCTTACCTGCGCGCTG 1440  
Db 1789 GGCAGCTCTACGGCGCGCGCCCTCTTTCTTCTGCTGCCAGCCTCTTACCTGCGCGCTG 1848  
QY 1441 CTCTGGCGCAAGTCTGTGGCCACCAAGTCTTCTTGCAGCTCGCCATCTTTCGTCATCGGC 1500  
Db 1849 CTCTGGCGCAAGTCTGTGGCCACCAAGTCTTCTTGCAGCTCGCCATCTTTCGTCATCGGC 1908  
QY 1501 GGCATCTGACGCTGTCCGCTTTCGTCATCTTCTCGAGCTCGCCATCTTTCGTCATCGGC 1560  
Db 1909 GGCATCTGACGCTGTCCGCTTTCGTCATCTTCTCGAGCTCGCCATCTTTCGTCATCGGC 1968  
QY 1561 ACCAAGCGGAGGACTAG 1578  
Db 1969 ACCAAGCGGAGGACTAG 1986  
RESULT 4  
ID ABK92051/C  
XX ABK92051 standard; DNA; 1734 BP.  
AC ABK92051;  
XX  
DT  
XX  
XX 14-AUG-2002 (first entry)  
XX  
DE DNA encoding novel GABA transporter-like receptor protein #3.  
XX  
KW Cancer; metabolic disorder; skin disorder; infectious disease; anorexia;  
KW behavioral disorder; valve disease; endocrine disorder; heart disorder;  
KW blood disorder; anxiety disorder; brain disorder; inflammatory disorder;  
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
KW immune disorder; haematopoietic disorder; dyslipidaemia;  
KW metabolic disturbance; obesity; metabolic syndrome X; wasting disorder;  
KW gene therapy; transgenic animal; human; gene; ds.  
XX  
OS Homo sapiens.  
XX  
XX WO200229058-A2.  
XX  
PD 11-APR-2002.  
XX  
XX 05-OCT-2001; 2001WO-US031248.  
XX  
XX 05-OCT-2000; 2000US-0238323P.  
PR 05-OCT-2000; 2000US-0238325P.  
PR 06-OCT-2000; 2000US-0238372P.  
PR 06-OCT-2000; 2000US-0238373P.  
PR 06-OCT-2000; 2000US-0238379P.

PR 06-OCT-2000; 2000US-0238382P.  
PR 06-OCT-2000; 2000US-0238383P.  
PR 06-OCT-2000; 2000US-0238384P.  
PR 06-OCT-2000; 2000US-0238397P.  
PR 06-OCT-2000; 2000US-0238400P.  
PR 06-OCT-2000; 2000US-0238401P.  
PR 06-OCT-2000; 2000US-0238402P.  
PR 14-MAR-2001; 2001US-0275892P.  
PR 08-JUN-2001; 2001US-0296860P.  
XX  
XX (CURA-) CURAGEN CORP.  
XX  
XX Shimkets RA, Taupier RJ, Burgess CE, Zerhusen BD, Mezes PS;  
PI Rastelli L, Malyankar UM, Grosse WM, Alsobrook JP, Lopley DM;  
PI Spytek KA, Li L, Edinger S, Gerlach V, Ellerman K, Macdougall J;  
PI Gunther E, Millet I, Stone D, Smithson G, Szekeres ES;  
XX  
XX WPI; 2002-444103/47.  
DR  
DR P-PSDB; ABG61784.  
XX  
PT Novel isolated polypeptide, designated NOVX, useful for treating or  
PT preventing cancer, diabetes, obesity, dyslipidemia, anorexia, and  
PT metabolic, neurodegenerative, immune and hematopoietic disorders.  
XX  
XX Claim 9; Page 62; 316pp; English.  
XX  
XX The invention describes an isolated polypeptide (I), useful in the  
CC manufacture of a medicament for treating a syndrome associated with a  
CC human disease. (I), the polynucleotide encoding it (II) and an antibody  
CC (III) to (I) are useful for treating or preventing cancer, metabolic  
CC disorders, skin disorders, infectious disease, anorexia, behavioral  
CC disorders, valve diseases, endocrine disorders, heart and blood  
CC disorders, anxiety disorders, brain disorders, inflammatory disorders,  
CC neurodegenerative disorder, Alzheimer's disease, Parkinson's disorder,  
CC dyslipidaemias, metabolic disturbances associated with obesity, and the various  
CC metabolic syndrome X and wasting disorders associated with chronic  
CC diseases. (I), (II) or (III) are useful in screening assays, detection  
CC assays (e.g., chromosomal mapping, tissue typing, forensic biology),  
CC predictive medicine (e.g., diagnostic assays and prognostic assays), and  
CC in methods of treatment (e.g., therapeutic and prophylactic). (I) is  
CC useful as immunogen to produce antibodies immunospecific for (I), to  
CC screen for potential agonist and antagonist compounds, and as bait  
CC protein in a two-hybrid or three-hybrid assay. (II) is useful in gene  
CC therapy, to express (I), to detect (I) mRNA or a genetic lesion in a (I)  
CC gene, and to modulate activity of (I). A cell containing a vector  
CC expressing (I) is useful for producing non-human transgenic animals. This  
CC sequence encodes a novel human polypeptide described in the invention  
XX  
XX Sequence 1734 BP; 358 A; 516 C; 578 G; 282 T; 0 U; 0 Other;  
SQ  
Query Match 99.9%; Score 1576.4; DB 6; Length 1734;  
Best Local Similarity 99.9%; Pred. No. 3.3e-311;  
Matches 1577; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 ATGGCCACCTTGTCTCGCAGCAGCTGTCCAGCTGGCCACGTCCTGTCCACAGTCC 60  
Db 1623 ATGGCCACCTTGTCTCGCAGCAGCTGTCCAGCTGGCCACGTCCTGTCCACAGTCC 1564  
QY 61 CAGGCCAAGATGAGCGGCATGTTCCGCCAGATGGGTTTTCAGCGCGCCACGGATGAGGAG 120  
Db 1563 CAGGCCAAGATGAGCGGCATGTTCCGCCAGATGGGTTTTCAGCGCGCCACGGATGAGGAG 1504  
QY 121 GCGGTGGGCTTCGCGCATTTGCCAGCACCTTCGACTTTTTCAGCCAGGCGCTTCGAGATG 180  
Db 1503 GCGGTGGGCTTCGCGCATTTGCCAGCACCTTCGACTTTTTCAGCGCGCCACGGATGAGGAG 1444  
QY 181 GACATCCTGAAAGCCGAGGAGAGCCCTCGCGGAGACGAGGCGCTGAAGCGCCGCTCGAG 240  
Db 1443 GACATCCTGAAAGCCGAGGAGAGCCCTCGCGGAGACGAGGCGCTGAAGCGCCGCTCGAG 1384  
QY 241 GGAGACATCATTTATCAGCGAGGCGAGCGAGTCTCTTCCCGCCCTCCGCTCCAAAGGAC 300



Db 1394 GGCAGCCTCAGCGGCGCGCTCTGTTCTTGTGCTCCAGCCTCTTTCACCTGCGCCTG 1453  
Qy 1441 CTCTGGCGAAGCTGCTGGACCAAGTCTTCTTCGAGTGGCATCTTCTGTCATCGGC 1500  
Db 1454 CTCTGGCGAAGCTGCTGGACCAAGTCTTCTTCGACGTGCCATCTTCTGTCATCGC 1513  
Qy 1501 GGCATCTGCAGCGTGCCTGCTTCTGTCACCTCCCTCGAGGGCTCATCGAAGCCTACCGA 1560  
Db 1514 GGCATCTGCAGCGTGCCTGCTTCTGTCACCTCCCTCGAGGGCTCATCGAAGCCTACCGA 1573  
Qy 1561 ACCAAGCGGAGGACTAG 1578  
Db 1574 ACCAAGCGGAGGACTAG 1591

RESULT 3  
ABK12349  
ID ABK12349 standard; DNA; 2261 BP.  
XX  
AC ABK12349;  
XX  
DT 18-JUN-2002 (first entry)  
XX  
DE Human gamma-aminobutyric acid (GABA) transporter DNA sequence.  
XX  
KW Human; gamma-aminobutyric acid transporter; GABA transporter; ds;  
KW Gene therapy; Alzheimer's disease; Parkinson's disease; epilepsy.  
XX  
OS Homo sapiens.  
XX  
PN WO200220753-A2.  
XX  
PD 14-MAR-2002.  
XX  
PF 28-AUG-2001; 2001WO-US026778.  
XX  
PR 01-SEP-2000; 2000US-0230178P.  
XX  
PA (LEXI-) LEXICON GENETICS INC.  
XX  
PI Friddle CJ, Gerhardt B, Hu Y;  
XX WPI; 2002-292265/33.  
XX  
DR New human nucleic acid, useful e.g. in drug screening, encodes gamma-  
PT aminobutyric acid transporter protein.  
XX  
PS Disclosure; Page 36; 36pp; English.

CC The present invention relates to a new nucleic acid that encodes a fully  
CC defined 525 amino acid sequence, as given in the specification, or  
CC hybridises under stringent conditions to a fully defined 1578 base pair  
CC sequence, or its complement, as given in the specification. The invention  
CC is useful for genomic mapping or identifying and validating exon splice  
CC junctions, in molecular mutagenesis and evolution of proteins, for  
CC recombinant production of encoded proteins, for production of transgenic  
CC knock-out or knock-in animals useful for identifying phenotypic  
CC expression of the nucleic acid of the invention and for assigning  
CC function to unknown genes, and also for generating antibodies to  
CC homologous or orthologous proteins. Other uses of the invention include  
CC as primers/probes for cloning, sequencing and library screening (for drug  
CC discovery), to characterise temporal and tissue-specific expression of  
CC genes, for identifying mutations and as a source of antisense, double-  
CC stranded or ribozyme reagents for gene therapy. The molecules of the  
CC invention and their fragments are used to raise antibodies, to identify  
CC related gene products, to screen for potential therapeutic agents, e.g.  
CC for treatment of Alzheimer's or Parkinson's diseases or epilepsy, i.e.  
CC (antagonists or other modulators), and as therapeutic agents. The present  
CC nucleic acid sequence represents the human gamma-aminobutyric acid (GABA)  
CC transporter DNA sequence of the invention. Human GABA transporters share  
CC structural similarity with mammalian GABA transporters

SQ Sequence 2261 BP; 348 A; 781 C; 680 G; 452 T; 0 U; 0 Other;  
Query Match 100.0%; Score 1578; DB 6; Length 2261;  
Best Local Similarity 100.0%; Pred. No. 1.ee-311;  
Matches 1578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ATGGCCACCTTGTCTCCGCGAGCAAGCTGTCCAAAGTGGCCAGTCCGTGTCCAAAGTCC 60  
Db 409 ATGGCCACCTTGTCTCCGCGAGCAAGCTGTCCAAAGTGGCCAGTCCGTGTCCAAAGTCC 468  
Qy 61 CAGGCCAAGATGAGCGGCATGTTCGCCAGGATGGGTTTTCAGGGCGCCACGATAGGAG 120  
Db 469 CAGGCCAAGATGAGCGGCATGTTCGCCAGGATGGGTTTTCAGGGCGCCACGATAGGAG 528  
Qy 121 CGGTGGGCTTTCGCGCATTTGCGACCTTGAGACCTTGAGACCGCCAGGGGCTGCAGATG 180  
Db 529 CGGTGGGCTTTCGCGCATTTGCGACCTTGAGACCTTGAGACCGCCAGGGGCTGCAGATG 588  
Qy 181 GACATCTGAAAGCGGAGGAGAGCCCTGCGGGGACGAGGGCGCTGAAAGCGCCCTCGAG 240  
Db 589 GACATCTGAAAGCGGAGGAGAGCCCTGCGGGGACGAGGGCGCTGAAAGCGCCCTCGAG 648  
Qy 241 GGAGACATCCATTTACGCGAGGAGCGGAGTCTCTTCCGCCCTCTCGGGTCCAAAGGAC 300  
Db 649 GGAGACATCCATTTACGCGAGGAGCGGAGTCTCTTCCGCCCTCTCGGGTCCAAAGGAC 708  
Qy 301 CAGGTGGGAGTGGTGGGCAATTCCGGGGCCACGACAGCCCAAAATCACGGCGTGGGAG 360  
Db 709 CAGGTGGGAGTGGTGGGCAATTCCGGGGCCACGACAGCCCAAAATCACGGCGTGGGAG 768  
Qy 361 GCAGGCTGGAACGTGACCAAGCCATCCAGGGCATGTTCTGCTGGGCTTACCTTACGCC 420  
Db 769 GCAGGCTGGAACGTGACCAAGCCATCCAGGGCATGTTCTGCTGGGCTTACCTTACGCC 828  
Qy 421 ATCTGACGCGCGCTACCTGGGGTGTCTTCTATCATCTTTCGCCCGTGTGTGTCTGC 480  
Db 829 ATCTGACGCGCGCTACCTGGGGTGTCTTCTATCATCTTTCGCCCGTGTGTGTCTGC 888  
Qy 481 TACACCGGCAAGATCTCATCGCGTGTCTGACGAGGAGATGAGAGCGGCGAGTGTGTG 540  
Db 889 TACACCGGCAAGATCTCATCGCGTGTCTGACGAGGAGATGAGAGCGGCGAGTGTGTG 948  
Qy 541 CGCGTGGCGGACTCTGTACGTGCCATAGCCAAAGCTGTCTGCCCGCCCGCTTCCCAAG 600  
Db 949 CGCGTGGCGGACTCTGTACGTGCCATAGCCAAAGCTGTCTGCCCGCCCGCTTCCCAAG 1008  
Qy 601 CTGGGCGCCGAGTGTGAACGTAGCGAGATCATTCGAGCTGTGTATGACGTGCATCTCTG 660  
Db 1009 CTGGGCGCCGAGTGTGAACGTAGCGAGATCATTCGAGCTGTGTATGACGTGCATCTCTG 1068  
Qy 661 TACGTGGTGTGAGTGGCAACCTCATGTATCAACAGCTTCCCGGGGCTGCCGTGTCCGAG 720  
Db 1069 TACGTGGTGTGAGTGGCAACCTCATGTATCAACAGCTTCCCGGGGCTGCCGTGTCCGAG 1128  
Qy 721 AAGTCTCTGTCATTTATCCCAAGCGCGCTGTCTGCTTGGCGCTTCTTAAAGAACCTC 780  
Db 1129 AAGTCTCTGTCATTTATCCCAAGCGCGCTGTCTGCTTGGCGCTTCTTAAAGAACCTC 1188  
Qy 781 AAGCGCGTGTCCAAGTTCAGTCTGTGTGCACTCTGGGCCACTTCTGTCATCAATATCTG 840  
Db 1189 AAGCGCGTGTCCAAGTTCAGTCTGTGTGCACTCTGGGCCACTTCTGTCATCAATATCTG 1248  
Qy 841 GTCATAGCTACTGTCTATCGGGGCGCGGAGTGGGCTGGGAGAGGTCAAGTCTTAC 900  
Db 1249 GTCATAGCTACTGTCTATCGGGGCGCGGAGTGGGCTGGGAGAGGTCAAGTCTTAC 1308  
Qy 901 ATCGAGCTCAAGAAGTTCCTCCATTTGGCATCTCTGTTTTCAGCTTACACGTCTCAG 960  
Db 1309 ATCGAGCTCAAGAAGTTCCTCCATTTGGCATCTCTGTTTTCAGCTTACACGTCTCAG 1368  
Qy 961 ATCTTCTGCGCTTCGCTGGAGGCAATATGACAGCAGCCAGCGAGTTCACCTGCATGATG 1020  
Db 1369 ATCTTCTGCGCTTCGCTGGAGGCAATATGACAGCAGCCAGCGAGTTCACCTGCATGATG 1428

KW candidate nucleus; amygdala; thalamus; psychiatric disorder; anxiety;  
KW neurological disorder; neurodegenerative disorder; bipolar disorder;  
KW unipolar disorder; schizophrenia; psychotic disorder; ss.  
XX Homo sapiens.  
OS  
XX  
XX  
XX Key Location/Qualifiers  
FH 14.1591  
CDS /\*tag= a  
FT /product= "neurotransmitter polypeptide HIPHUM 0000057"  
FT  
XX  
XX GB2366566-A.  
XX  
XX  
XX 13-MAR-2002.  
XX  
XX 31-MAY-2001; 2001GB-00013257.  
XX  
XX 01-JUN-2000; 2000GB-00013239.  
XX  
XX (GLAX ) GLAXO GROUP LTD.  
XX  
XX Sala CF, Terstappen GC;  
XX  
XX WPI; 2002-397154/43.  
XX P-PSDB; ABB09876.  
XX  
XX Isolated neurotransmitter polypeptide HIPHUM 0000057, useful as a drug  
XX target for treating neurological/neurodegenerative disorders, e.g.  
XX bipolar disorder, unipolar disorder, anxiety, schizophrenia and/or  
XX psychotic disorders.  
XX  
XX Claim 5; Page 23-26; 35pp; English.  
XX  
XX The present sequence encodes a human neurotransmitter polypeptide,  
XX designated HIPHUM 0000057. HIPHUM 0000057 is mapped to chromosomal region  
XX 20q12-q13. HIPHUM 0000057 is specifically expressed in the brain, and the  
XX highest level of expression is in the candidate nucleus. High levels of  
XX expression are also found in the amygdala and thalamus. These areas of the  
XX brain are associated with psychiatric, neurological and neurodegenerative  
XX disorders. The neurotransmitter transporter polypeptide HIPHUM 0000057 is  
XX a screening target for identification and development of drugs, i.e.  
XX modulators of neurotransmitter transporter polypeptide expression and/or  
XX activity. These drugs may be used in the prevention and/or treatment of  
XX psychiatric disorders or neurological/neurodegenerative disorders, e.g.  
XX bipolar disorder, unipolar disorder, anxiety, schizophrenia and/or  
XX psychotic disorders  
XX  
XX SQ Sequence 1595 BP; 268 A; 521 C; 480 G; 326 T; 0 U; 0 Other;  
Query Match 100.0%; Score 1578; DB 6; Length 1595;  
Best Local Similarity 100.0%; Pred. No. 1.5e-311;  
Matches 1578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 ATGGCCACCTTGCTCCGACGAAAGTGTCCAAAGTGGCCACGTCCTGTGTCACCAAGTCC 60  
14 ATGGCCACCTTGCTCCGACGAAAGTGTCCAAAGTGGCCACGTCCTGTGTCACCAAGTCC 73  
61 CAGGCCAAGATAGCGGCATGTTCCGACGATGGGTTTTCAGCGCGCCACGATAGGAG 120  
74 CAGGCCAAGATAGCGGCATGTTCCGACGATGGGTTTTCAGCGCGCCACGATAGGAG 133  
121 GCGGTGGGCTTCGCGCATTCGCGACGACCTCGATTGAGCACCGCCAGGCGCTCGAGATG 180  
134 GCGGTGGGCTTCGCGCATTCGCGACGACCTCGATTGAGCACCGCCAGGCGCTCGAGATG 193  
181 GACATCCTGAAAGCCGAGGAGAGCCCTCGCGGAGACGAGGCGCTGAAAGCGCCGTCGAG 240  
194 GACATCCTGAAAGCCGAGGAGAGCCCTCGCGGAGACGAGGCGCTGAAAGCGCCGTCGAG 253  
241 GGAGACATTCATTATCAGCGAGGAGCGGAGTCTCTTGCCTCCGCTCCGCTCCAAAGGAC 300  
254 GGAGACATTCATTATCAGCGAGGAGCGGAGTCTCTTGCCTCCGCTCCGCTCCAAAGGAC 313

QY 301 CAGGTGGAGGTGGTGGCGAATTCGGGGGCCACGACAAAGCCCAAAATACAGCGGTGGAG 360  
DB 314 CAGGTGGAGGTGGTGGCGAATTCGGGGGCCACGACAAAGCCCAAAATACAGCGGTGGAG 373  
QY 361 GCAGGCTGGAAACGTGACCAAACGCGCATCCAGGGCATGTTCTGTGTGGGCTTACCCCTACGCC 420  
DB 374 GCAGGCTGGAAACGTGACCAAACGCGCATCCAGGGCATGTTCTGTGTGGGCTTACCCCTACGCC 433  
QY 421 ATCTCTGACCGCGGCTACCTGGGGTGTGTTCTCATCATCTTGGCGCGGTGTGTGTGTC 480  
DB 434 ATCTCTGACCGCGGCTACCTGGGGTGTGTTCTCATCATCTTGGCGCGGTGTGTGTGTC 493  
QY 481 TACACGGCAAGATCTCTCATCGGTGCTGTACGAGGAGAAATGAAGACGCGAGGTGGTG 540  
DB 494 TACACGGCAAGATCTCTCATCGGTGCTGTGTACGAGGAGAAATGAAGACGCGAGGTGGTG 553  
QY 541 CGCGTGGGACCTCGTACGTGGCCATAGCAACGCGCTGTGCGGCCCGCGCTTCCCAACG 600  
DB 554 CGCGTGGGACCTCGTACGTGGCCATAGCAACGCGCTGTGCGGCCCGCGCTTCCCAACG 613  
QY 601 CTGGGCGCGCGAGTGGTGAACGTAGCGCAGATCATCGAGCTGTGTATGACGTGATCCTG 660  
DB 614 CTGGGCGCGCGAGTGGTGAACGTAGCGCAGATCATCGAGCTGTGTATGACGTGATCCTG 673  
QY 661 TACGTGGTGGTGAAGTGGCAACCTCATGTACAAAGTGTCCCGGGGCTGCGGTGTCGAG 720  
DB 674 TACGTGGTGGTGAAGTGGCAACCTCATGTACAAAGTGTCCCGGGGCTGCGGTGTCGAG 733  
QY 721 AAGTCTCTGTCATTATCGCACGCGCGTGTCTGCTGCTTGGCGCTTCTTAAAGAACCTC 780  
DB 734 AAGTCTCTGTCATTATCGCACGCGCGTGTCTGCTGCTTGGCGCTTCTTAAAGAACCTC 793  
QY 781 AAGCGCGTGTCAAGTTCAGTCTGTGTGACATCTCTGGCCCACTTCTGTATCAATATCCTG 840  
DB 794 AAGCGCGTGTCAAGTTCAGTCTGTGTGACATCTCTGGCCCACTTCTGTATCAATATCCTG 853  
QY 841 GTCATAGCTTACTGTCTATCGCGGGCGCGAGCTGGGCTTGGGAGAGGTCAAGTTCATC 900  
DB 854 GTCATAGCTTACTGTCTATCGCGGGCGCGAGCTGGGCTTGGGAGAGGTCAAGTTCATC 913  
QY 901 ATCGAGCTCAAGAAAGTTCCTCCATCTCATTTGGCATCATGTGTTCAGTACACGTCTCAG 960  
DB 914 ATCGAGCTCAAGAAAGTTCCTCCATCTCATTTGGCATCATGTGTTCAGTACACGTCTCAG 973  
QY 961 ATCTTCTGCTTTCGTGTGAGGCAATATGACAGCCAGCGAGTTCATCGATGATG 1020  
DB 974 ATCTTCTGCTTTCGTGTGAGGCAATATGACAGCCAGCGAGTTCATCGATGATG 1033  
QY 1021 AACTGACGCACATCGCAGCTGCGTGTCTCAAGGGGCTCTTTCGCGCTCGTTCGCTACCTC 1080  
DB 1034 AACTGACGCACATCGCAGCTGCGTGTCTCAAGGGGCTCTTTCGCGCTCGTTCGCTACCTC 1093  
QY 1081 ACCTGGCGCGACGAGACCAAGGAGGTTCATACGGATAAATCTGCGCGGCTCCATCCCGGCC 1140  
DB 1094 ACCTGGCGCGACGAGACCAAGGAGGTTCATACGGATAAATCTGCGCGGCTCCATCCCGGCC 1153  
QY 1141 GTGGTCAACATCTTCTGTGTGCGCAAGGCGTGTGTTCTCTATCTCTGCGCATTTCTTGGC 1200  
DB 1154 GTGGTCAACATCTTCTGTGTGCGCAAGGCGTGTGTTCTCTATCTCTGCGCATTTCTTGGC 1213  
QY 1201 GCTGTGAGGTGCTGGAGAGTCTCTTTCAGGAGGAGCGCGCGCTTTTTCGCGGCC 1260  
DB 1214 GCTGTGAGGTGCTGGAGAGTCTCTTTCAGGAGGAGCGCGCGCTTTTTCGCGGCC 1273  
QY 1261 TGCTACAGCGCGGACGCGGCGCTTGAAGTCTTGGGGCTGACCGTGGCGCTGCGCGCTCGTC 1320  
DB 1274 TGCTACAGCGCGGACGCGGCGCTTGAAGTCTTGGGGCTGACCGTGGCGCTGCGCGCTCGTC 1333  
QY 1321 GTCTTCAAGCTGTCTATGCGCAATTTATGTGCGGCACTTTCGCGCTGTCTATGCGGCTCAACC 1380  
DB 1334 GTCTTCAAGCTGTCTATGCGCAATTTATGTGCGGCACTTTCGCGCTGTCTATGCGGCTCAACC 1393  
QY 1381 GCGAGCTTCAAGCGGCGCGGCTCTGTGTTTCTGCTGCGCCAGCGCTCTTTTCACTGCGGCT 1440

CC sequence, or its complement, as given in the specification. The invention  
CC is useful for genomic mapping or identifying and validating exon splice  
CC junctions, in molecular mutagenesis and evolution of proteins, for  
CC recombinant production of encoded proteins, for production of transgenic  
CC knock-out or knock-in animals useful for identifying phenotypic  
CC expression of the nucleic acid of the invention and for assigning  
CC function to unknown genes, and also for generating antibodies to  
CC homologous or orthologous proteins. Other uses of the invention include  
CC as primers/probes for cloning, sequencing and library screening (for drug  
CC discovery), to characterise temporal and tissue-specific expression of  
CC genes, for identifying mutations and as a source of antisense, double-  
CC stranded or ribozyme reagents for gene therapy. The molecules of the  
CC invention and their fragments are used to raise antibodies, to identify  
CC related gene products, to screen for potential therapeutic agents, e.g.  
CC for treatment of Alzheimer's or Parkinson's diseases or epilepsy, i.e.  
CC (ant)agonists or other modulators, and as therapeutic agents. The present  
CC nucleic acid sequence encodes the human gamma-aminobutyric acid (GABA)  
CC transporter protein of the invention. The protein shares structural  
CC similarity with mammalian GABA transporters

SQ Sequence 1578 BP; 267 A; 510 C; 475 G; 326 T; 0 U; 0 Other;

Query Match 100.0%; Score 1578; DB 6; Length 1578;

Best Local Similarity 100.0%; Pred. No. 1.5e-311;

Matches 1578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGGCCACCTTCTCGCAGCAGCTGCCAAGTGGCCACCTCGTGTCCAAAGTCC	60
DB	1	ATGGCCACCTTCTCGCAGCAGCTGCCAAGTGGCCACCTCGTGTCCAAAGTCC	60
QY	61	CAGGCCAAGATGAGCGGATGTTCCGACGATGGTTCAGCGCGCCACCGATGAGG	120
DB	61	CAGGCCAAGATGAGCGGATGTTCCGACGATGGTTCAGCGCGCCACCGATGAGG	120
QY	121	GGGTGGGCTTCGCGATTTGCGACGACCTTCGATGAGCAGCGAGGCGCTGCGATG	180
DB	121	GGGTGGGCTTCGCGATTTGCGACGACCTTCGATGAGCAGCGAGGCGCTGCGATG	180
QY	181	GACATCTCAAGACCGAGGAGAGCCCTCGGGGAGAGGGCGCTGAGAGCGCCGTCGAG	240
DB	181	GACATCTCAAGACCGAGGAGAGCCCTCGGGGAGAGGGCGCTGAGAGCGCCGTCGAG	240
QY	241	GGAGACATCATTTATCAGCGAGGAGCGAGCTCTCTCGCCGCTTCGCGCTCCAAAGGAC	300
DB	241	GGAGACATCATTTATCAGCGAGGAGCGAGCTCTCTCGCCGCTTCGCGCTCCAAAGGAC	300
QY	301	CAGGTGGAGGTGGTGGCGAATTCGGGGGCGACGACAGCCCAAAATCAGCGCGTGGAG	360
DB	301	CAGGTGGAGGTGGTGGCGAATTCGGGGGCGACGACAGCCCAAAATCAGCGCGTGGAG	360
QY	361	GCAGGCTGGAACGTGACCAACGCCATCCAGGCGATGTTCTGCTGGGCTACCCCTACGCC	420
DB	361	GCAGGCTGGAACGTGACCAACGCCATCCAGGCGATGTTCTGCTGGGCTACCCCTACGCC	420
QY	421	ATCTGTCAGCGGGCTTACCTGGGGTGTTCATCATCTTCGCGCGCGTGTGTGCTGC	480
DB	421	ATCTGTCAGCGGGCTTACCTGGGGTGTTCATCATCTTCGCGCGCGTGTGTGCTGC	480
QY	481	TACACCGCAAGATCCTCATTCGCGTCTGTGACGAGGAGATGAAGACGCGAGGTGGT	540
DB	481	TACACCGCAAGATCCTCATTCGCGTCTGTGACGAGGAGATGAAGACGCGAGGTGGT	540
QY	541	CGCGTGGGACTCGTACGCGCATAGCAAGCGCTGCTGGCGCCCGGCTTCCCAAGC	600
DB	541	CGCGTGGGACTCGTACGCGCATAGCAAGCGCTGCTGGCGCCCGGCTTCCCAAGC	600
QY	601	CTGGGCGGCGAGTGGTGAACGTAGCGCAGATCATCGAGCTGGTATGACGTGCTCTG	660
DB	601	CTGGGCGGCGAGTGGTGAACGTAGCGCAGATCATCGAGCTGGTATGACGTGCTCTG	660
QY	661	TACGTGGTGGTGGAGGCACTCATGTATCAACAGCTTCCCGGGCTGCCGTGCGCAG	720
DB	661	TACGTGGTGGTGGAGGCACTCATGTATCAACAGCTTCCCGGGCTGCCGTGCGCAG	720

QY	721	AAGTCTCGTCCATTATCGCCACGCGCGTGTGCTGCTGCTGCGCTTCTTTAAGAACCTC	780
DB	721	AAGTCTCGTCCATTATCGCCACGCGCGTGTGCTGCTGCTGCGCTTCTTTAAGAACCTC	780
QY	781	AAGCGCGGTGCCAAGTTCAGTCTGTGTGCACTCTGGCCCACTTCGTCAATCAATATCTG	840
DB	781	AAGCGCGGTGCCAAGTTCAGTCTGTGTGCACTCTGGCCCACTTCGTCAATCAATATCTG	840
QY	841	GTCTAGCTACTGTCTATCCGGGCGCGGACCTGGGCTGGGAGAGGTCAAGTTCATC	900
DB	841	GTCTAGCTACTGTCTATCCGGGCGCGGACCTGGGCTGGGAGAGGTCAAGTTCATC	900
QY	901	ATCGACGTCAAGAAGTTCCTCCATTTGGCATCATCGTGTTCAGCTACACGTCTCAG	960
DB	901	ATCGACGTCAAGAAGTTCCTCCATTTGGCATCATCGTGTTCAGCTACACGTCTCAG	960
QY	961	ATCTTCTGCTTCTCGTGGAGGCAATATGAGCAGAGCCAGCGAGTTCACCTGCATGATG	1020
DB	961	ATCTTCTGCTTCTCGTGGAGGCAATATGAGCAGAGCCAGCGAGTTCACCTGCATGATG	1020
QY	1021	AAGTGGACGACATCGCAGCGCTGCTCAAGGCGCTTTCGCGCTGCTGCGCTACCTC	1080
DB	1021	AAGTGGACGACATCGCAGCGCTGCTCAAGGCGCTTTCGCGCTGCTGCGCTACCTC	1080
QY	1081	ACCTGGGCGGACGAGACCAAGGAGGTATCACGGATAACCTGCCCGGCTCCATCCGCGCC	1140
DB	1081	ACCTGGGCGGACGAGACCAAGGAGGTATCACGGATAACCTGCCCGGCTCCATCCGCGCC	1140
QY	1141	GTGTGCAACATCTTTCTGGTGGCCAAAGCGGTGTTCTATCTCTGCTGCTTCTTGGC	1200
DB	1141	GTGTGCAACATCTTTCTGGTGGCCAAAGCGGTGTTCTATCTCTGCTGCTTCTTGGC	1200
QY	1201	GCTGTGAGGTGCTGGAGAGTCCGCTTCCAGGAGGAGCGCGCTTTTCCCGGCC	1260
DB	1201	GCTGTGAGGTGCTGGAGAGTCCGCTTCCAGGAGGAGCGCGCTTTTCCCGGCC	1260
QY	1261	TGCTACAGCGCGACGCGCGCTGAAGTCTTGGGGGCTGAGCGCTGCGCTCGCTCGCTC	1320
DB	1261	TGCTACAGCGCGACGCGCGCTGAAGTCTTGGGGGCTGAGCGCTGCGCTCGCTCGCTC	1320
QY	1321	GTCTTACGCTGCTCATGGCCATTTATGTGCGCAGCTTTCGGGCTGCTCATGGGCTCACC	1380
DB	1321	GTCTTACGCTGCTCATGGCCATTTATGTGCGCAGCTTTCGGGCTGCTCATGGGCTCACC	1380
QY	1381	GGCAGCTTCAAGGCGCGCGCTCTGTTTCTGCTGCGCGAGCTTTCACCTGCGCGCTG	1440
DB	1381	GGCAGCTTCAAGGCGCGCGCTCTGTTTCTGCTGCGCGAGCTTTCACCTGCGCGCTG	1440
QY	1441	CTCTGGCGCAAGCTGCTGTGGCACCAGTCTTCTTCGACGTGCGCATCTTTCGTCTATCGGC	1500
DB	1441	CTCTGGCGCAAGCTGCTGTGGCACCAGTCTTCTTCGACGTGCGCATCTTTCGTCTATCGGC	1500
QY	1501	GGCATCTCAGCGTTCGGCTTCGTGCACTCCCTTCGAGGGGCTCATCGAAGCTACCGA	1560
DB	1501	GGCATCTCAGCGTTCGGCTTCGTGCACTCCCTTCGAGGGGCTCATCGAAGCTACCGA	1560
QY	1561	ACCAACGCGGAGGACTAG 1578	
DB	1561	ACCAACGCGGAGGACTAG 1578	

## RESULT 2

ABL56565

ID ABL56565 standard; cDNA; 1595 BP.

XX ABL56565;

XX AC

XX 05-AUG-2002 (first entry)

XX DT

DE Nucleotide sequence of neurotransmitter polypeptide HIPHUM 0000057.

XX Human; gene; neurotransmitter; HIPHUM 0000057; 20q12-q13; brain;

XX KW

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: November 3, 2004, 18:26:25 ; Search time 750 Seconds  
(without alignments)  
11044.782 Million cell updates/sec

Title: US-09-940-919-1  
Perfect score: 1578  
Sequence: 1 atggccacctgtctccgag.....gaaccaacggaggactag 1578

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 4134886 seqs, 2624710521 residues  
Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq\_23Sep04: \*  
1: geneseqn1980s: \*  
2: geneseqn1980s: \*  
3: geneseqn2000s: \*  
4: geneseqn2001as: \*  
5: geneseqn2001bs: \*  
6: geneseqn2002as: \*  
7: geneseqn2002bs: \*  
8: geneseqn2003as: \*  
9: geneseqn2003bs: \*  
10: geneseqn2003cs: \*  
11: geneseqn2003ds: \*  
12: geneseqn2004s: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1578	100.0	1578	6 ABK12348	Abk12348 cDNA enco
2	1578	100.0	1595	6 ABL56565	Ab156565 Nucleotid
3	1578	100.0	2261	6 ABK12349	Abk12349 Human gam
C 4	1576.4	99.9	1734	6 ABK92051	Abk92051 DNA encod
5	1576.4	99.9	1763	6 ABK92049	Abk92049 DNA encod
6	1576.4	99.9	1800	5 AAS15704	Aas15704 Human ves
C 7	1574.8	99.8	1734	6 ABK92050	Abk92050 DNA encod
C 8	1574.8	99.8	1734	6 ABK92052	Abk92052 DNA encod
C 9	1574.8	99.8	1734	6 ABK92053	Abk92053 DNA encod
C 10	1574.8	99.8	1734	6 ABK92054	Abk92054 DNA encod
11	1344.2	85.2	1648	5 AAS06061	Aas06061 Angiotens
C 12	1140.2	72.2	1723	5 AAS06087	Aas06087 Angiotens
13	387.2	24.5	1095	5 AAS81399	Aas81399 DNA encod
14	366	23.2	1650	4 ABL14045	Ab114045 Drosophil
15	365	23.2	3722	4 ABL14044	Ab114044 Drosophil
16	237	15.0	255	9 ACH15349	Ach15349 Human adu
17	232.2	14.7	2740	12 ADQ08701	Adq08701 Ciona int
18	177.2	11.2	498	9 ACH38839	Ach38839 Human foe
19	131	8.3	448	5 ABA20103	Aba20103 Human ner
20	101.8	6.5	451	5 ABA13679	Aba13679 Human ner
21	77.2	4.9	1580	8 ADA70336	Ada70336 Rice gene

22	72.6	4.6	1269	8 ADA70222	Ada70222 Rice gene
23	70.4	4.5	1257	8 ADA69963	Ada69963 Rice gene
24	65	4.1	1821	10 ADK41684	Adk41684 Maize ami
25	64.6	4.1	1212	12 ADJ39267	Adj39267 Plant CDW
26	62.2	3.9	2000	8 ADA71938	Ada71938 Rice gene
C 27	60.6	3.8	1050	4 AAD21685	Aad21685 Mutationa
C 28	60.6	3.8	2849	4 AAD21684	Aad21684 Human ret
C 29	59.8	3.8	109519	5 AAS08693	Aas08693 Micromono
30	58	3.7	1134	8 ADA71193	Ada71193 Rice gene
31	58	3.7	1260	9 ADA48445	Ada48445 Rice gene
C 32	58	3.7	3489	3 AAA30290	Aaa30290 Kaposi's
C 33	58	3.7	3489	4 AAF82901	Aaf82901 Nucleotid
C 34	58	3.7	3489	6 ABA93487	Ab93487 Kaposi's
C 35	58	3.7	3489	12 ADJ65095	Adj65095 HIV8 DNA
36	58	3.7	32207	2 AAV73805	Aav73805 KSHV LUR
37	58	3.7	137507	2 AAV19941	Aav19941 KSHV long
38	58	3.7	137508	12 ADN12162	Adn12162 Human her
39	57.2	3.6	1996	12 ADM47942	Adm47942 Polynucle
40	57	3.6	536	10 ADB68842	Adb68842 Minoroty
41	57	3.6	90597	10 ADJ72363	Adj72363 Streptomy
42	57	3.6	90600	6 ABQ78872	Abq78872 S. rooseo
C 43	56.4	3.6	628	12 ACH89664	Ach89664 Human gen
44	56	3.5	1700	10 ADK41690	Adk41690 Maize ami
C 45	55.6	3.5	2000	8 ADA71938	Ada71938 Rice gene

ALIGNMENTS

RESULT 1  
ABK12348  
ID ABK12348 standard; cDNA; 1578 BP.  
XX  
AC ABK12348;  
XX  
DT 18-JUN-2002 (first entry)  
XX  
DE cDNA encoding human gamma-aminobutyric acid (GABA) transporter protein.  
XX  
KW Human; gamma-aminobutyric acid transporter; GABA transporter; gene; ss;  
KW gene therapy; Alzheimer's disease; Parkinson's disease; epilepsy.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..1578  
FT /\*tag= a  
FT /product= "Human gamma-aminobutyric acid (GABA)  
FT transporter protein"  
XX  
FN WO200220753-A2.  
XX  
PD 14-MAR-2002;  
XX  
PF 28-AUG-2001; 2001WO-US026778.  
XX  
PR 01-SEP-2000; 2000US-0230178P.  
XX  
PA (LEXI-) LEXICON GENETICS INC.  
XX  
PI Friddle CJ, Gerhardt B, Hu Y;  
XX  
DR WPI; 2002-292265/33.  
DR P-PSDB; AAU78281.  
XX  
PT New human nucleic acid, useful e.g. in drug screening, encodes gamma-aminobutyric acid transporter protein.  
XX  
PS Claim 1; Page 34; 36pp; English.  
XX  
CC The present invention relates to a new nucleic acid that encodes a fully defined 525 amino acid sequence, as given in the specification, or hybridises under stringent conditions to a fully defined 1578 base pair

```
Qy 421 ATCTGCGAGCGGCTACTCGGGTTGTTCTCATCATCTTCGCGCGGTTGTGTGCTGC 480
Db 421 ATCTGCGAGCGGGCTACTCGGGTTGTTCTCATCATCTTCGCGCGGTTGTGTGCTGC 480
Qy 481 TACACCGGCAAGATCCTCATCGCTGCTGTACGAGGAGAAATGAAGCGGAGGTGGTG 540
Db 481 TACACCGGCAAGATCCTCATCGCTGCTGTACGAGGAGAAATGAAGCGGAGGTGGTG 540
Qy 541 CGCGTGGGACTCTGTAGTGCCATAGCCAAAGCTGTGCGGCCCGCGGCTTCCCAAG 600
Db 541 CGCGTGGGACTCTGTAGTGCCATAGCCAAAGCTGTGCGGCCCGCGGCTTCCCAAG 600
Qy 601 CTGGGCGGCGAGTGTGAACGTAGCGAGATCATCGAGCTGTGATGACGTGCATCTTG 660
Db 601 CTGGGCGGCGAGTGTGAACGTAGCGAGATCATCGAGCTGTGATGACGTGCATCTTG 660
Qy 661 TACGTGGTGGTGAAGTGGCAACCTCATGTACAACAGCTTCCCGGGCTGCCCGTGTGCGAG 720
Db 661 TACGTGGT----- 668
Qy 721 AAGTCTGTGTCATTATCGCCACGCGCGTGTGCTGCGCTTCCCTTAAGAACCTC 780
Db 669 -----GAACCTC 675
Qy 781 AAGGCGGTGTCGAAGTTCAGTCTGTGCTGTCACCTCTGGCCACTTCGTGATCAATATCCTG 840
Db 676 AAGGCGGTGTCGAAGTTCAGTCTGTGCTGTCACCTCTGGCCACTTCGTGATCAATATCCTG 735
Qy 841 GTCATAGCCTACTGTCTATCGCGGCGCGGAGCTGGGCTGGGAGAGGTCAAGTTCTAC 900
Db 736 GTCATAGCCTACTGTCTATCGCGGCGCGGAGCTGGGCTGGGAGAGGTCAAGTTCTAC 795
Qy 901 ATCGAGCTCAAGAAGTTCGCCATCTCCATTGGCATCATGTTTCAGCTACACGTCTCAG 960
Db 796 ATCGAGCTCAAGAAGTTCGCCATCTCCATTGGCATCATGTTTCAGCTACACGTCTCAG 855
Qy 961 ATCTTCTGCTTCTGCTGGAGGGCAATATGCAGAGCCAGCGAGTTCACCTGCAATGATG 1020
Db 856 ATCTTCTGCTTCTGCTGGAGGGCAATATGCAGAGCCAGCGAGTTCACCTGCAATGATG 915
Qy 1021 AACTGGACGCACATCGAGCCTCGTGTCTCAAGGGCCTCTTCGCGCTCGTGCCTACCTC 1080
Db 916 AACTGGACGCACATCGAGCCTCGTGTCTCAAGGGCCTCTTCGCGCTCGTGCCTACCTC 975
Qy 1081 ACCTGGCGCGAGAGCAAGAGGTATACGGATAACCTGCCGGCTCCATCCGCGCC 1140
Db 976 ACCTGGCGCGAGAGCAAGAGGTATACGGATAACCTGCCGGCTCCATCCGCGCC 1035
Qy 1141 GTGGTCAACATCTTCTGTGGCCAAAGCGCTGTTGCTATCTCTGCAATCTTTTGGC 1200
Db 1036 GTGGTCAACATCTTCTGTGGCCAAAGCGCTGTTGCTATCTCTGCAATCTTTTGGC 1095
Qy 1201 GCTGTGAGGTGTGGAGAAGTCTCTTCCAGGAAGGAGCGCGGCTTTTCCCGGCC 1260
Db 1096 GCTGTGAGGTGTGGAGAAGTCTCTTCCAGGAAGGAGCGCGGCTTTTCCCGGCC 1155
Qy 1261 TCGTACAGCGGCGAGCGGGCGCTGAAGTCTGCGGGCTGACGCTGGGCTGCGCGCTCGTC 1320
Db 1156 TCGTACAGCGGCGAGCGGGCGCTGAAGTCTGCGGGCTGACGCTGGGCTGCGCGCTCGTC 1215
Qy 1321 GTCTTCACGCTGTCTATGGCCATTTATGTGCCACTTCGCGCTGTCTATGGGCTCACCC 1380
Db 1216 GTCTTCACGCTGTCTATGGCCATTTATGTGCCACTTCGCGCTGTCTATGGGCTCACCC 1275
Qy 1381 GGCAGCCTCACGGGCGCGGCTCTGTCTTCTGCTGCCAGCTCTTTTCACTGCGCCTG 1440
Db 1276 GGCAGCCTCACGGGCGCGGCTCTGTCTTCTGCTGCCAGCTCTTTTCACTGCGCCTG 1335
Qy 1441 CTCTGGCGCAAGTGTGTGGACCAAGTCTTCTTCGAGCTGCGCATCTTGTCTATCGGC 1500
Db 1336 CTCTGGCGCAAGTGTGTGGACCAAGTCTTCTTTCGAGCTGCGCATCTTGTCTATCGGC 1395
```

```
Qy 1501 GGCATCTGACGCTGTCCGGCTTCGTCACCTCCCTCGAGGGCTCATCGAAGCCTTACCGA 1560
Db 1396 GGCATCTGACGCTGTCCGGCTTCGTCACCTCCCTCGAGGGCTCATCGAAGCCTTACCGA 1455
Qy 1561 ACCAACGCGGAGGACTAG 1578
Db 1456 ACCAACGCGGAGGACTAG 1473
```

Search completed: November 3, 2004, 21:05:54  
Job time : 6616 secs

QY	271	GCTCTCTTGGCGCCTCTCGGCTCCAAAGGACAGGTGGAGGTGTGGCGAAATTCGGGGGC	330
DB	121	GCTCTCTTGGCGCCTCTCGGCTCCAAAGGACAGGTGGAGGTGTGGCGAAATTCGGGGGC	180
QY	331	CACGACAAGCCCAAATCACCGCGTGGGAGCAGCGCTGGAACGTGACCAAGCCATCCAG	390
DB	181	CACGACAAGCCCAAATCACCGCGTGGGAGCAGCGCTGGAACGTGACCAAGCCATCCAG	240
QY	391	GGCATGTTGTCTGGGCTTACCCTACGCCATCTGCAACGCGCTACTCTGGGCTGTGTT	450
DB	241	GGCATGTTGTCTGGGCTTACCCTACGCCATCTGCAACGCGCTACTCTGGGCTGTGTT	300
QY	451	CTCATATCTTTGGCGCGTTGTGTGTCTTACACCGCAAGATCCTCATCGCTGCGCTG	510
DB	301	CTCATATCTTTGGCGCGTTGTGTGTCTTACACCGCAAGATCCTCATCGCTGCGCTG	360
QY	511	TACGAGGAGATGAAGACGGCGAGTGGTGGCGTGGGGGACTGTTAGTGGCCATAGCC	570
DB	361	TACGAGGAGATGAAGACGGCGAGTGGTGGCGTGGGGGACTGTTAGTGGCCATAGCC	420
QY	571	AACGCTCTCTCGCGCCCGCTTCCAAACGCTGGGCGCGGAGTGTGAACGTAGCGCAG	630
DB	421	AACGCTCTCTCGCGCCCGCTTCCAAACGCTGGGCGCGGAGTGTGAACGTAGCGCAG	480
QY	631	ATCATTCGAGCTGGTGAAGCTGATCCTGTACGTGGTGGTGGTGGGCAACCTCATGTAC	690
DB	481	ATCATTCGAGCTGGTGAAGCTGATCCTGTACGTGGTGGTGGTGGGCAACCTCATGTAC	540
QY	691	AACAGCTTCCGGGGCTGCCCGTGTGGAGAAGTCTCTGGTCCATATTCGCCACGGCCGTG	750
DB	541	AACAGCTTCCGGGGCTGCCCGTGTGGAGAAGTCTCTGGTCCATATTCGCCACGGCCGTG	600
QY	751	CTGCTGCCCTTGGCGCTTCTTAAAGACCTCAAGCGCGTGTCCAAAGTTCAGTCTGCTGC	810
DB	601	CTGCTGCCCTTGGCGCTTCTTAAAGACCTCAAGCGCGTGTCCAAAGTTCAGTCTGCTGC	660
QY	811	ACTCTGGGCCACTTCTGTCATCAATATCTGGTCTATAGCTACTGCTTATCGCGGGCGCG	870
DB	661	ACTCTGGGCCACTTCTGTCATCAATATCTGGTCTATAGCTACTGCTTATCGCGGGCGCG	720
QY	871	GACTGGGCTGGGAGAGGTCAAGTTCTATACGACGTCAAGAAAGTTCGCCATCTCCATT	930
DB	721	GACTGGGCTGGGAGAGGTCAAGTTCTATACGACGTCAAGAAAGTTCGCCATCTCCATT	780
QY	931	GGCATCATGTGTTACAGTACACGCTCAGATCTTCTGCTTCGCTGGAGGGCAATATG	990
DB	781	GGCATCATGTGTTACAGTACACGCTCAGATCTTCTGCTTCGCTGGAGGGCAATATG	840
QY	991	CAGCAGCCACGAGTTCCACTGCATGATGAACGTGACGCACTCGCAGCGCTGGTGCTC	1050
DB	841	CAGCAGCCACGAGTTCCACTGCATGATGAACGTGACGCACTCGCAGCGCTGGTGCTC	900
QY	1051	AAGGGCTTTTGGCGCTGTCGCTTCACTCACTTGGGCCGACGAGACCAAGGAGTCAATC	1110
DB	901	AAGGGCTTTTGGCGCTGTCGCTTCACTCACTTGGGCCGACGAGACCAAGGAGTCAATC	960
QY	1111	ACGATTAACCTGCCCGGCTCCATCCGGCGCGTGTCAACATCTTTCTGGTGGCCAGGGG	1170
DB	961	ACGATTAACCTGCCCGGCTCCATCCGGCGCGTGTCAACATCTTTCTGGTGGCCAGGGG	1020
QY	1171	CTGTTGTCTTACTCTCTGCCATTCTTTTGGCGCTGTCGAGTGTGGAGAAAGTGCCTTTC	1230
DB	1021	CTGTTGTCTTACTCTCTGCCATTCTTTTGGCGCTGTCGAGTGTGGAGAAAGTGCCTTTC	1080
QY	1231	CAGGAAGCAGCGCGCTTTTTCGCCCTGTCTACAGCGCGACGCGCGCTCGAAGTCC	1290
DB	1081	CAGGAAGCAGCGCGCTTTTTCGCCCTGTCTACAGCGCGACGCGCGCTCGAAGTCC	1140
QY	1291	TGGGGGTGACGTGGCGTGGCGCTGTGTCTTTCACGCTGTCTATGGCCATTTATGTG	1350
DB	1141	TGGGGGTGACGTGGCGTGGCGCTGTGTCTTTCACGCTGTCTATGGCCATTTATGTG	1200

Qy	1351	CCGCACTTCGGCGCTGCTCATGAGGCTTACCGGAGAGCTTACAGGGCGCGGCTCTGTGTTTC	1411
Db	1201	CCGCACTTCGGCGCTGCTCATGAGGCTTACCGGAGAGCTTACAGGGCGCGGCTCTGTGTTTC	1260
Qy	1411	TTGCTGCGCCAGCCCTTTTACCTTCGGCGCTGCTGCGGCAAGCTGTGTGGCACCAAGTC	1470
Db	1261	TTGCTGCGCCAGCCCTTTTACCTTCGGCGCTGCTGCGGCAAGCTGTGTGGCACCAAGTC	1320
Qy	1471	TTCTTTGACGCTGCGCATCTTTCGTCATCGGCGGCATCTGACAGCGTGTCCGGCTTCGTGCAC	1530
Db	1321	TTCTTTGACGCTGCGCATCTTTCGTCATCGGCGGCATCTGACAGCGTGTCCGGCTTCGTGCAC	1380
Qy	1531	TCCCTCGAGGGCTTCATCGAAGCTTACCGAACCAACCGCGAGGACTAG	1578
Db	1381	TCCCTCGAGGGCTTCATCGAAGCTTACCGAACCAACCGCGAGGACTAG	1428
RESULT 15			
CQ727095		1473 bp	DNA
LOCUS		Sequence 13029 from Patent WO02068579.	linear
DEFINITION			PAT 03-FEB-2004
ACCESSION			
VERSION		CQ727095.1	GI:42292118
KEYWORDS			
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
TITLE		Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.	
		Kits, such as nucleic acid arrays, comprising a majority of	
		humanexons or transcripts, for detecting expression and other uses	
		thereof	
JOURNAL		Patent: WO 02068579-A 13029 06-SEP-2002;	
FEATURES		PE Corporation (NY) (US)	
source		Location/Qualifiers	
		1..1473	
		/organism="Homo sapiens"	
		/mol_type="unassigned DNA"	
		/db_xref="taxon:9606"	
ORIGIN			
Query Match		86.1%;	Score 1358; DB 6; Length 1473;
Best Local Similarity		93.3%;	Pred. No. 3.3e-196;
Matches 1473; Conservative		0; Mismatches	0; Indels 105; Gaps 1;
Qy	1	ATGGCCACCTTGCTCCGACGAAAGCTGTCAACGTGGCCACGTCGTGTGCCAACAAAGTC	60
Db	1	ATGGCCACCTTGCTCCGACGAAAGCTGTCAACGTGGCCACGTCGTGTGCCAACAAAGTC	60
Qy	61	CAGGCCAAGATGAGCGGCATGTTTCGCCAGGATGGGTTTTTCAGGCGGCCACGGATGAGGAG	120
Db	61	CAGGCCAAGATGAGCGGCATGTTTCGCCAGGATGGGTTTTTCAGGCGGCCACGGATGAGGAG	120
Qy	121	CGCGTGGGCTTCGCGCATTTGCGACGACCTCGACTTTTGAGCACCGCCAGGCGCTTCGACAGATG	180
Db	121	CGCGTGGGCTTCGCGCATTTGCGACGACCTCGACTTTTGAGCACCGCCAGGCGCTTCGACAGATG	180
Qy	181	GACATCCTGAAAGCCGAGGGAGAGCCCTCGCGGGACGAGGGCGCTGGAAGCGCCGCTCGAG	240
Db	181	GACATCCTGAAAGCCGAGGGAGAGCCCTCGCGGGACGAGGGCGCTGGAAGCGCCGCTCGAG	240
Qy	241	GGAGACATCCATTATCAGCGAGCGAGGAGTCTCTCTCGCGGCTTCGCGCTTCGAGGAC	300
Db	241	GGAGACATCCATTATCAGCGAGCGAGGAGTCTCTCTCGCGGCTTCGCGCTTCGAGGAC	300
Qy	301	CAGGTGGAGGTGGTGGCGAATTCGGGGGCGACGACAAAGCCCAAATATCACGGCGTGGGAG	360
Db	301	CAGGTGGAGGTGGTGGCGAATTCGGGGGCGACGACAAAGCCCAAATATCACGGCGTGGGAG	360
Qy	361	GCAGGCTGGAAACGTGACCAACGCCATTCAGGGGCGTTCGTCGTGGGCTTACCTACGCC	420
Db	361	GCAGGCTGGAAACGTGACCAACGCCATTCAGGGGCGTTCGTCGTGGGCTTACCTACGCC	420

1576 GTGGTCAACATCTTTCTGGTGGCCAAAGGGCTGTGTCTCTATCTCTCTGCAATCTTTCGCC 1635  
QY 1201 GCTGTCGAGTCTGAGAGTCTCTCTTCCAGGAGGACGCGCGCTTTTCCCGGCC 1260  
Db 1636 GCTGTCGAGTCTGAGAGTCTCTTCCAGGAGGACGCGCGCTTTTCCCGGCC 1695  
QY 1261 TGCTACAGCGGACGCGCGCTGAAGTCTTGGGGGCTGAGCTGCGTGGCGCTCGTC 1320  
Db 1696 TGCTACAGCGGACGCGCGCTGAAGTCTTGGGGGCTGAGCTGCGTGGCGCTCGTC 1755  
QY 1321 GTCTTACGCTGCTCATGCGCATTTATGTGCGCGCATCTTGGCGCTGCTCATGCGCTCAC 1380  
Db 1756 GTCTTACGCTGCTCATGCGCATTTATGTGCGCGCATTTGCGCTGCTCATGCGCTCAC 1815  
QY 1381 GGCAGCCTCAGCGGCGCGCTGTCTTCTGCTGCCAGCTCTTTCACCTGCGCTG 1440  
Db 1816 GGCAGCCTCAGCGGCGCGCTGTCTTCTGCTGCCAGCTCTTTCACCTGCGCTG 1875  
QY 1441 CTCTGGCGCAAGCTGCTGTGGCACCAGTCTTCTTTCGAGCTGCGCATCTTGTCTCATCGGC 1500  
Db 1876 CTCTGGCGCAAGCTGCTGTGGCACCAGTCTTCTTTCGAGCTGCGCATCTTGTCTCATCGGC 1935  
QY 1501 GGCATCTGCGAGTCTGCGCTGCTGTCGATCTTCCCTCGAGGCGCTCATGAGGCTTACCGA 1560  
Db 1936 GGCATCTGCGAGTCTGCGCTGCTGTCGATCTTCCCTCGAGGCGCTCATGAGGCTTACCGA 1995  
QY 1561 ACCACGCGGAGGACTAG 1578  
Db 1996 ACCACGCGGAGGACTAG 2013

RESULT 14  
BC036458 2174 bp mRNA linear PRI 30-SEP-2003  
LOCUS Homo sapiens vesicular inhibitory amino acid transporter, mRNA  
DEFINITION (cDNA clone IMAGE:4813726), partial cds.  
ACCESSION BC036458  
VERSION  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM  
REFERENCE 1 (bases 1 to 2174)  
AUTHORS Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L.,  
Scheetz T.E., Brownstein M.J., Ustin T.B., Toshiyuki S.,  
Carninci P., Frange C., Raha S.S., Loquellano N.A., Peters G.J.,  
Abramson R.D., Mullaly S.J., Bosak S.A., McEwan P.J.,  
McKernan K.J., Malek J.A., Gunaratne P.H., Richards S.,  
Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
Fahy J., Helton E., Kettner M., Madan A., Rodriguez S.,  
Sanchez A., Whiting M., Young A.C., Touchman J.W., Green E.D.,  
Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D.,  
Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
Butterfield Y.S., Krzywinski M.I., Skalska U., Smalish D.E.,  
Schnarch A., Schein J.E., Jones S.J. and Marra M.A.  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
22388257 12477932  
2 (bases 1 to 2174)  
Strausberg R.  
Direct Submission  
Submitted (09-AUG-2002) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

REMARK  
COMMENT

USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
On Sep 16, 2003 this sequence version replaced gi:22209088.  
Contact: MGC help desk  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki  
Toshiyuki and Piero Carninci (RIKEN)  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Institute for Systems Biology  
<http://www.systemsbio.org>  
contact: [amadanesystemsbiology.org](mailto:amadanesystemsbiology.org)  
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha  
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAK Plate: 46 Row: d Column: 1.

## FEATURES

## source

1. 2174  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4813726"  
/tissue\_type="Brain, hippocampus"  
/clone\_lib="NIH MGC\_95"  
/lab\_host="DH10B"  
/note="Vector: pBluescript"  
<1. 2174  
/gene="VIAAT"  
/note="synonym: ba12201.1"  
/db\_xref="LocusID:140679"  
<1. 1428  
/gene="VIAAT"  
/codon\_start=1  
/product="VIAAT protein"  
/protein\_id="AAH36458.2"  
/db\_xref="GI:34785016"  
/db\_xref="LocusID:140679"  
/translation="GFSEHROGLMDILKAECEPCGEGAEAPVEGDIHVORGSGAPLP  
PSGKDQVGGGEGFHDKPKITAEAGNVTNAIQNFVLGPAALHGGVGLGLFI  
IFAAVCCYTGKILIALENEDEGVVRDVSVAIANACAPFPFLLGRGVNVAQ  
IELVMTCLIVVYVGNLYNSFPGLPVSKWSIIATAVLLPCAFKLNLRVSKFSL  
LCTLAHFVNLIVAYICLSRADWAEKVFYIDVKFPIGIGIVFYSTQIFPLSL  
EGNQDSEPHCMNWTTHAACVLKGLFALVYLTVADTKVDITDNLPGSTRVAVNI  
FLVAKLLVPLPFAAVLEKSLFQSGRAFPACVSGDGRKLSGLTLRCALVVF  
TLMAIYVPHFALMLGTGSLTGAGLCFLFLSLFHLRLKLLMHWQVFPDVAIFVG  
GICSVSGFVHSLGLELIEAYRTNAED"  
193. 1398  
/note="VIAAT"  
/gene="VIAAT"; Region: Transmembrane amino acid  
transporter protein. This transmembrane region is found in  
many amino acid transporters including UNC-47 and MTR.  
UNC-47 encodes a vesicular amino butyric acid (GABA)  
transporter, (VGAT). MTR is a N system amino acid  
transmembrane domains. MTR is a N system amino acid  
transporter system protein involved in methyltryptophan  
resistance. Other members of this family include proline  
transporters and amino acid permeases"  
/db\_xref="CDD:pfam01490"

## gene

## CDS

## misc\_feature

## ORIGIN

Query Match 90.3%; Score 1424.8; DB 9; Length 2174;  
Best Local Similarity 99.9%; Pred. No. 2.4e-206;  
Matches 1426; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 151 GACTTTGAGCACCGCAGGCGCTGAGATGACATCTCTGAAAGCCGAGGAGCGCTGC 210  
Db 1 GGCCTTTGAGCACCGCAGGCGCTGAGATGACATCTCTGAAAGCCGAGGAGCGCTGC 60  
QY 211 GGGGACGAGGCGCTGAAGCGCCCTCGAGGAGACATCCATTATCAGGAGCGACCGGA 270  
Db 61 GGGGACGAGGCGCTGAAGCGCCCTCGAGGAGACATCCATTATCAGGAGCGACCGGA 120



# KEYWORDS SOURCE ORGANISM

oligo capping; fis (full insert sequence).  
Macaca fascicularis (crab-eating macaque)  
Macaca fascicularis  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
Cercopithecinae; Macaca.

## REFERENCE AUTHORS

Oada, N., Hida, M., Kusuda, J., Tanuma, R., Iseki, K., Hirai, M.,  
Terao, K., Suzuki, Y., Sugano, S. and Hashimoto, K.  
Isolation of full-length cDNA clones from macaque brain cDNA  
libraries

## TITLE

Unpublished  
2 (bases 1 to 2756)  
Hashimoto, K., Oada, N., Hida, M., Kusuda, J. and Sugano, S.  
Direct Submission

## JOURNAL REFERENCE AUTHORS

Submitted (11-JUN-2001) Katsuyuki Hashimoto, National Institute of  
Infectious Diseases, Division of Genetic Resources; 23-1, Toyama  
1-chome, Shinjuku-ku, Tokyo 162-8640, Japan  
(E-mail: khashi@nih.go.jp, URL: http://www.nih.go.jp/yoken/genebank/,  
Tel: 81-3-5285-1111 (ex. 2120), Fax: 81-3-5285-1181)

## COMMENT

Lab host: TOP10  
Vector: pME18S-FL3 (Acc.No. AB009864)  
R. Site1: DraIII (CACTGTGTG)  
R. Site2: DraIII (CACCATGTG)  
Description: 1st strand cDNA was primed with an oligo (dT) primer  
[ATGCGCTTTTCTTTTCTTTT]; double-stranded cDNA was synthesized  
using specific 5' and 3' primers and amplified by PCR. The PCR  
product was digested with SfiI and size selection was performed to  
exclude fragments <1.5kb. The SfiI-digested PCR product was cloned  
into distinct DraIII sites of pME18S-FL3. XhoI sites just outside  
the DraIII sites can be used to isolate the cDNA insert. Libraries  
were constructed by oligo-capping method  
(Sugano et al., , Institute of Medical Science, University of  
Tokyo).

Custom primer used for sequencing  
( 5' end primer [CTTCTGCTCTAAAGCTGG];  
3' end primer [CGACCTGCAGTCCGAGCAC]).

## FEATURES source

Location/Qualifiers  
1..2756  
/organism="Macaca fascicularis"  
/mol\_type="mRNA"  
/db\_xref="taxon:9541"  
/clone="OccE-21148"  
/sex="male"  
/tissue\_type="cerebellum cortex"  
/clone\_lib="macaque brain cDNA library QccB"  
/dev\_stage="adult"  
436..2013  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="BAB60726.1"  
/db\_xref="GI:114388326"

## CDS

496 CAGGCCAAGGTGAGCGCATGTTCCGACAGATGGTTTTCAGGGCGCCACGGATGAGGAG 555  
121 GCGGTGGGCTTCGCGCATTCGACGACCTTCGACTTTGAGCACCACGAGGCGCTGCAAGATG 180  
556 GCGGTGGGCTTCGCGCATTCGACGACCTTCGACTTTGAGCACCACGAGGCGCTGCAAGATG 615  
181 GACATCTGAAAGCCGAGGAGAGCCCTGCGGGAGAGAGCGGCTGAGAGCGCGCTGCGAG 240  
616 GACATCTGAAAGCTGAGGCGAGCGCTGCGGGAGAGAGCGGCTGAGAGCGCGCTGCGAG 675  
241 GGAGACATCCATTATCAGCGAGGAGCGGAGCTCTTGCCTGCGGCTCCGCTCCAAAGGAC 300  
676 GGAGACATCCATTATCAGCGGCGGAGCGGCTCTTGCCTGCGGCTCCGCTCCAAAGGAC 735  
301 CAGGTGGGAGTGGTGGCGAATTCGGGGGCCACGACAAAGCCCAAAATACAGGCGTGGAG 360  
736 CAGGTGGGAGTGGTGGTGGTGGGCGGCGACGACAAAGCCCAAGATACAGGCGTGGAG 795  
361 GCAGGCTGGAACGTGACCAACGCGCATCCAGGCGATGTTCTGCTGGGCTTACCTACGCC 420  
796 GCAGGCTGGAACGTGACCAACGCGCATCCAGGCGATGTTCTGCTGGGCTTACCTACGCC 855  
421 ATCTCTGACGCGGCTTACCTGCGGTTGTTTCTCATCATCTTCCGCGCGTGTGTGCTGC 480  
856 ATCTCTGACGCGGCTTACCTGCGGTTGTTTCTCATCATCTTCCGCGCGTGTGTGCTGC 915  
481 TACACGCGAAGATCTCATTCGCGTCTGTACGAGGAGAAATGAGACGCGAGGTGGT 540  
916 TACACTGGCAAGATCTCATTCGCGTCTGTACGAGGAGAAATGAGACGCGAGGTGGT 975  
541 CCGGTGGGACTCTGCTGAGCGCATAGCCAGCGCTGCTGCGCGCGGCTTCCCAAGC 600  
976 CCGGTGGGACTCTTACGTGGCATAGCCAGCGCTGCTGCGCGCGGCTTCCCAAGC 1035  
601 CTGGCGCGCGAGTGGTGAACGTAGCGAGATCATCGAGCTGGTGTGATGAGCTGCATCTC 660  
1036 CTGGCGCGCGAGTGGTGAACGTGGCGAGATCATCGAGCTGGTGTGATGAGCTGCATCTC 1095  
661 TAGCTGGTGGTGGGCAACCTCATGTACAACAGCTTCCCGGGCTGCGCGTGCAG 720  
1096 TAGCTGGTGGTGGGCAACCTCATGTACAACAGCTTCCCGGGCTGCGCGTGCAG 1155  
721 AGTCTGTGTCATTCGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780  
1156 AGTCTGTGTCATTCGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1215  
781 AAGGCGCTGTCCAAGTTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840  
1216 AAGGCGCTGTCCAAGTTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1275  
841 GTCATAGCTACTGTCTATTCGCGGGCGCGAGCTGGGCTGGGAGAGGTCAGAGTTCATC 900  
1276 GTCATAGCTACTGTCTATTCGCGGGCGCGAGCTGGGCTGGGAGAGGTCAGAGTTCATC 1335  
901 ATCGAGCTCAAGAGTTCCCATCTCCATGGCATCATGTGTTCAGCTACAGTCTCAG 960  
1336 ATCGAGCTCAAGAGTTCCCATCTCCATGGCATCATGTGTTCAGCTACAGTCTCAG 1395  
961 ATCTTCTGCTTCTGCTGGAGGCAATATGACAGCGCCAGCGAGTTCACACTGCATGATG 1020  
1396 ATCTTCTGCTTCTGCTGGAGGCAATATGACAGCGCCAGCGAGTTCACACTGCATGATG 1455  
1021 AACTGGACGACATCGACGCTGCGGTCTCAAGGGCTCTTCCGCGCTGCTGCGCTACCTC 1080  
1456 AACTGGACGACATCGCGCTGCGGTCTCAAGGGCTCTTCCGCACTGCTGCGCTACCTC 1515  
1081 ACCTGGCGCGAGAGACGAGGTCATCAACGATAAATGCGCGGCTCCATCCGGGC 1140  
1516 ACCTGGCGCGAGAGACGAGGTCATCAACGATAAATGCGCGGCTCCATCCGGGC 1575  
1141 GTGGTCAACATCTTTCTGCTGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200

## ORIGIN

Query Match 96.2%; Score 1518.8; DB 9; Length 2756;  
Best Local Similarity 97.7%; Pred. No. 1.3e-220;  
Matches 1541; Conservative 0; Mismatches 37; Indels 0; Gaps 0;  
QY 1 ATGGCCACCTTGCTCGCAGCAAGCTGTCCAAGCTGGCGACGCTCGCTGTCCAAAGTCC 60  
DB 436 ATGGCCACCTTGCTCGCAGCAAGCTGTCCAAGCTGGCGACGCTCGCTGTCCAAAGTCC 495  
QY 61 CAGGCCAAGATGAGCGCATGTTCCGCGAGGATGGGTTTTCAGGCGGCCAGGATGAGGAG 120



LOCUS AX497266 1733 bp DNA linear PAT 26-SEP-2002  
DEFINITION Sequence 33 from Patent WO0229058.  
ACCESSION AX497266  
VERSION AX497266.1 GI:23342642

## KEYWORDS

Homo sapiens (human)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 Shinketsu, R.A., Taupier, R.J., Burgess, C.E., Zerhusen, B.D.,  
Mezes, P.S., Rastelli, L., Malyankar, U.M., Grosse, W.M.,  
Alsbrook, J.P., Lepley, D.M., Spytek, K.A., Li, L., Edinger, S.,  
Gerlach, V., Ellemman, K., Macdougall, J., Gunther, E., Millet, I.,  
Stone, D., Smithson, G. and Szekeres, E.S.  
Human proteins, polynucleotides encoding them and methods of using  
the same

## TITLE

JOURNAL Patent: WO 0229058-A 33 11-APR-2002;  
Curagen Corporation (US)

## FEATURES

1..1733  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

## ORIGIN

Query Match 99.3%; Score 1567.4; DB 6; Length 1733;  
Best Local Similarity 99.9%; Pred. No. 6.2e-228;  
Matches 1568; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 ATGGCCACCTTCTCGCAGCAGAGTGTCCAACTGGCCACGTCCTGGTGTCCAAAGTCC 60  
DB 112 ATGGCCACCTTCTCGCAGCAGAGTGTCCAACTGGCCACGTCCTGGTGTCCAAAGTCC 171  
  
QY 61 CAGGCCAAGATGAGCGGATGTTTCCGAGGATGGGTTTTCAGCGGCGCACGATGAGGAG 120  
DB 172 CAGGCCAAGATGAGCGGATGTTTCCGAGGATGGGTTTTCAGCGGCGCACGATGAGGAG 231  
  
QY 121 GCGGTGGGCTTCGCGATTTGCGACGACCTCGACTTTTGAGCAGCGCCAGCGCTCGAGATG 180  
DB 232 GCGGTGGGCTTCGCGATTTGCGACGACCTCGACTTTTGAGCAGCGCCAGCGCTCGAGATG 291  
  
QY 181 GACATCTCTGAAGCCGAGGAGAGCCCTCGCGGAGCAGGGCGCTGAAGCGCCCTCGAG 240  
DB 292 GACATCTCTGAAGCCGAGGAGAGCCCTCGCGGAGCAGGGCGCTGAAGCGCCCTCGAG 351  
  
QY 241 GGAGACATCCATTATCAGCGAGCGAGCGAGTCTCTTCGCGCCCTTCGCGCTCGAAGGAC 300  
DB 352 GGAGACATCCATTATCAGCGAGCGAGCGAGTCTCTTCGCGCCCTTCGCGCTCGAAGGAC 411  
  
QY 301 CAGGTGGGAGGTGGTGGCGAATTCGGGGCCACGACCAAGCCAAATCAGCGCGTGGGAG 360  
DB 412 CAGGTGGGAGGTGGTGGCGAATTCGGGGCCACGACCAAGCCAAATCAGCGCGTGGGAG 471  
  
QY 361 GCAGGCTGGAACGTGACCAACGCCATCCAGGGCATGTTCTGCTGGGCGCTACCCCTACGCC 420  
DB 472 GCAGGCTGGAACGTGACCAACGCCATCCAGGGCATGTTCTGCTGGGCGCTACCCCTACGCC 531  
  
QY 421 ATCTCGACGGCGGCTTACCTGGGTTGTTTCTCATCATCTTCGCGCGCGTGTGTGCTGC 480  
DB 532 ATCTCGACGGCGGCTTACCTGGGTTGTTTCTCATCATCTTCGCGCGCGTGTGTGCTGC 591  
  
QY 481 TACACGGCAAGATCCTCATCGGTCCTGTACGAGGAATGAGACGGCGAGGTGGTG 540  
DB 592 TACACGGCAAGATCCTCATCGGTCCTGTACGAGGAATGAGACGGCGAGGTGGTG 651  
  
QY 541 CCGGTGGGAGTCTGTACGTCGTCATAGCCATAGCCCTGTGCGGCCCGCGCTTCCCAAGC 600  
DB 652 CCGGTGGGAGTCTGTACGTCGTCATAGCCATAGCCCTGTGCGGCCCGCGCTTCCCAAGC 711  
  
QY 601 CTGGGCGGCGAGTGTGTAACGTGAGCGAGATCATCGAGCTGTGTATGATGATGATCTCTG 660  
DB 712 CTGGGCGGCGAGTGTGTAACGTGAGCGAGATCATCGAGCTGTGTATGATGATGATCTCTG 771

QY 661 TACGTGGTGGTGGAGTGGCAACCTCATGTACAAACAGCTTCCCGGGCTGCCGTGTGCGAG 720  
DB 772 TACGTGGTGGTGGAGTGGCAACCTCATGTACAAACAGCTTCCCGGGCTGCCGTGTGCGAG 831  
  
QY 721 AAGTCTGGTCCATTATCGCCACGCGCGTGTGTGCTTGGCGCTTCCCTAAAGAACCTC 780  
DB 832 AAGTCTGGTCCATTATCGCCACGCGCGTGTGTGCTTGGCGCTTCCCTAAAGAACCTC 891  
  
QY 781 AAGCGGTGTCCTCAAGTTCAGTCTGCTGCACTCTGGCCCACTTCGTCTCATCAATATCTCG 840  
DB 892 AAGCGGTGTCCTCAAGTTCAGTCTGCTGCACTCTGGCCCACTTCGTCTCATCAATATCTCG 951  
  
QY 841 GTCATAGCTACTGTCTATCCGCGGCGCGCACTGGGCTTGGGAGAGGTCAAGTGTCTAC 900  
DB 952 GTCATAGCTACTGTCTATCCGCGGCGCGCACTGGGCTTGGGAGAGGTCAAGTGTCTAC 1011  
  
QY 901 ATCGACGTCGAAGAGTTCCTCCATCTCATTTGGCATCATCTGTGTTGAGTACACGTCCTCAG 960  
DB 1012 ATCGACGTCGAAGAGTTCCTCCATCTCCATTGGCATCATCTGTGTTGAGTACACGTCCTCAG 1071  
  
QY 961 ATCTTCTGCTTCCGCTGGAGGCAATATGAGCAGAGCCAGGAGTTCACCTGCAATGATG 1020  
DB 1072 ATCTTCTGCTTCCGCTGGAGGCAATATGAGCAGAGCCAGGAGTTCACCTGCAATGATG 1131  
  
QY 1021 AACTGGACGACATCGCAGCCTGCTCAAGGGCTCTTTCGCGCTCGTTCGCTTACCTC 1080  
DB 1132 AACTGGACGACATCGCAGCCTGCTCAAGGGCTCTTTCGCGCTCGTTCGCTTACCTC 1191  
  
QY 1081 ACCTGGGCGGACGAGACCAAGAGGTATACAGGATAACCTGCGCGGCTTCATCCGCGCC 1140  
DB 1192 ACCTGGGCGGACGAGACCAAGAGGTATACAGGATAACCTGCGCGGCTTCATCCGCGCC 1251  
  
QY 1141 GTGTCAACATCTTCTGCTGGTGGCAAGCGCTGTGCTCTCTCTGCTGCTCTTCTTGGC 1200  
DB 1252 GTGTCAACATCTTCTGCTGGTGGCAAGCGCTGTGCTCTCTCTGCTGCTCTTCTTGGC 1311  
  
QY 1201 GCTGTCAAGGTGCTGGAGAGTCTCTTTCAGGAAGGACGCGCGCTTTTTCGCGGCC 1260  
DB 1312 GCTGTCAAGGTGCTGGAGAGTCTCTTTCAGGAAGGACGCGCGCTTTTTCGCGGCC 1371  
  
QY 1261 TGCTACAGCGGCGAGCGGCGCTGAAGTCTCTGGGGGTGAGCGTGGCTGCGCGCTCGTC 1320  
DB 1372 TGCTACAGCGGCGAGCGGCGCTGAAGTCTCTGGGGGTGAGCGTGGCTGCGCGCTCGTC 1431  
  
QY 1321 GCTTTCAGCTGCTCATGCGCCATTTATGTGCGGCACTTTCGCGCTCATGCGCGCTCACC 1380  
DB 1432 GCTTTCAGCTGCTCATGCGCCATTTATGTGCGGCACTTTCGCGCTCATGCGCGCTCACC 1491  
  
QY 1381 GGCAGCTTCAGCGGCGCGGCTCTGTTTCTTGTGCGCCAGCTCTTTCACCTGCGCGCTG 1440  
DB 1492 GGCAGCTTCAGCGGCGCGGCTCTGTTTCTTGTGCGCCAGCTCTTTCACCTGCGCGCTG 1551  
  
QY 1441 CTCTGGCGCAAGTGTCTGTGCAACCAAGTCTTTCGAGCGTGCATCTTTCGTCATCGGC 1500  
DB 1552 CTCTGGCGCAAGTGTCTGTGCAACCAAGTCTTTCGAGCGTGCATCTTTCGTCATCGGC 1611  
  
QY 1501 GGCATCTCGAGGCTGCGGCTTGTGCACTCCCTCGAGGGCTCATCGAAGCCTACCGA 1560  
DB 1612 GGCATCTCGAGGCTGCGGCTTGTGCACTCCCTCGAGGGCTCATCGAAGCCTACCGA 1671  
  
QY 1561 ACCAACGCG 1569  
DB 1672 ACCAACGCG 1680

## RESULT 13

AB062931

## LOCUS

AB062931

## DEFINITION

Macaca fascicularis brain cDNA clone.QcseE-21148, full insert

## ACCESSION

AB062931

## VERSION

AB062931.1 GI:14388325

## AUTHORS

Shimkets, R.A., Taupier, R.J., Burgess, C.E., Zerkhusen, B.D.,  
Mezes, P.S., Rastelli, L., Malyankar, U.M., Grosse, W.M.,  
Alsobrook, J.P., Lepley, D.M., Spytek, K.A., Li, L., Edinger, S.,  
Gerlach, V., Ellerman, K., MacDougall, J., Gunther, E., Millet, I.,  
Stone, D., Smithson, G. and Szekeres, E.S.  
Human proteins, polynucleotides encoding them and methods of using  
the same

## JOURNAL

Patent: WO 0229058-A 39 11-APR-2002;  
Curagen Corporation (US)

## FEATURES

## source

## Location/Qualifiers

1..1734

/organism="Homo sapiens"

/mol\_type="unassigned DNA"

/db\_xref="taxon:9606"

112..1689

/note="unnamed protein product"

/codon\_start=1

/protein\_id="CAD48688.1"

/db\_xref="GI:23342648"

/translation="MATLLRSKLSNVATSVNSKSOAKSGMFARMGFOADTEAVGF  
AHCDDLFERHQGLQMDILKAEGPCGDEGAEPVEGDIHQKSGAPLPSPSGKDV  
GGGFEHDKPKITAWERAGWNTNATQGMVLGLPYAILHGLYGLFLIFAAVCC  
YTKILLIAEENEDDEVVRDYSVAIANACCAPFTLGGRVNVAQIIEIWTFC  
ILYVVSNGLMYNSFPGLPVSKWSIIATVLLPCAPFLKRLKAVSFSLCTLAHFV  
INILVIACLSRARDWAEKVFYIDVKEPISIGIIVFSYTSQILPLPSLEGNMOOS  
EFCMMWTHIAACVLKGLFALVAYLTWADETKEAIDNLPISRIVAVNIFLAKALL  
SYPLPFAAVELEKSLFQSGRAFFPACYSGGRLKWSGLTLCALVVFLLMAIYV  
PHFALLMGLTGTAGLCFLPLPSLFLRLWRLKLLWHQVFFDVAIFVIGGICSVSGF  
VHSLEGLIEAYRTNAED"

## ORIGIN

Query Match 99.8%; Score 1574.8; DB 6; Length 1734;  
Best Local Similarity 99.9%; Pred. No. 4.7e-229;  
Matches 1576; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGCCACCTTGTCCGAGCAAGCTGTCCAACTGGCCAGCTCCGTGTCCAAAGTCC 60  
DB 112 ATGCCACCTTGTCCGAGCAAGCTGTCCAACTGGCCAGCTCCGTGTCCAAAGTCC 171

QY 61 CAGGCCAAGATAGCGGCATGTTCCGCCAGATGGGTTTTCAGCGGCCACCGATGAGGAG 120  
DB 172 CAGGCCAAGATAGCGGCATGTTCCGCCAGATGGGTTTTCAGCGGCCACCGATGAGGAG 231

QY 121 CGCGTGGGCTTCGGCATTCGGACGACCTCGACTTTGAGCACCGCCAGGSCCTGCAGATG 180  
DB 232 CGCGTGGGCTTCGGCATTCGGACGACCTCGACTTTGAGCACCGCCAGGSCCTGCAGATG 291

QY 181 GACATCTGAAAGCCGAGGAGAGCCCTCGGGGACGAGCGCTGAAGCGCCCGTCGAG 240  
DB 292 GACATCTGAAAGCCGAGGAGAGCCCTCGGGGACGAGCGCTGAAGCGCCCGTCGAG 351

QY 241 GGAGACATCCATTATCAGCGAGGAGCGGAGCTCTCTGCGGCCCTCCGGCTCCAAGGAC 300  
DB 352 GGAGACATCCATTATCAGCGAGGAGCGGAGCTCTCTGCGGCCCTCCGGCTCCAAGGAC 411

QY 301 CAGGTGGGAGTGTGGCGAATTCGGGGGACACGACAAAGCCAAATACAGGCTGGGAG 360  
DB 412 CAGGTGGGAGTGTGGCGAATTCGGGGGACACGACAAAGCCAAATACAGGCTGGGAG 471

QY 361 CGAGCTTGAACGCTGACCAAGCCATCCAGGCAATGTTCTGCTGCGSCCTACCTACGCC 420  
DB 472 CGAGCTTGAACGCTGACCAAGCCATCCAGGCAATGTTCTGCTGCGSCCTACCTACGCC 531

QY 421 ATCTGACGCGCGCTACCTGGGTTGTTTCTCATCATCTTTCGCGCGCTGTGTGCTGC 480  
DB 532 ATCTGACGCGCGCTACCTGGGTTGTTTCTCATCATCTTTCGCGCGCTGTGTGCTGC 591

QY 481 TACACCGCAAGATCTCATCGCTGCTGTACGAGGAGATGAAGACGCGAGGTGGTG 540  
DB 592 TACACCGCAAGATCTCATCGCTGCTGTACGAGGAGATGAAGACGCGAGGTGGTG 651

QY 541 CGCGTGGGACTCGTACGTGGCCATAGCCAAAGCCCTGCTGCGCCCGCTTCCCAACG 600

DB 652 CGCGTGGGACTCGTACGTGGCCATAGCCAAAGCCCTGCTGCGCCCGCTTCCCAACG 711  
QY 601 CTGGGCGGCGAGTGTGTGAACGTAGCCAGATCATTCAGCTGTGTGATGAGCTGATCCTG 660  
DB 712 CTGGGCGGCGAGTGTGTGAACGTAGCCAGATCATTCAGCTGTGTGATGAGCTGATCCTG 771  
QY 661 TAGTGTGTGTGAGTGGCAACCTCATGTCAACAGACTTCCCGGGCTGCCGTGTGCGAG 720  
DB 772 TAGTGTGTGTGAGTGGCAACCTCATGTCAACAGACTTCCCGGGCTGCCGTGTGCGAG 831  
QY 721 AAGTCTGTGTCCATTATTCGCGACGCGCGTGTCTGCTGCTTTCGCGCTTCTTAAAGACCTC 780  
DB 832 AAGTCTGTGTCCATTATTCGCGACGCGCGTGTCTGCTGCTTTCGCGCTTCTTAAAGACCTC 891  
QY 781 AAGCCGTGTGCCAAGTTCAGTCTGTGTGCACTCTTGCCCACTTCTGTCATCAATATCCTG 840  
DB 892 AAGCCGTGTGCCAAGTTCAGTCTGTGTGCACTCTTGCCCACTTCTGTCATCAATATCCTG 951  
QY 841 GTCATAGCCTACTGTCTATCGCGGCGCGCATGCGGCTTGGGAGAGGTCAAGTTCATCTAC 900  
DB 952 GTCATAGCCTACTGTCTATCGCGGCGCGCATGCGGCTTGGGAGAGGTCAAGTTCATCTAC 1011  
QY 901 ATCGACGTCAAGAAAGTTCCTCCATCTCATTTGGCATCATCGTGTTCAGCTACACGTCTCAG 960  
DB 1012 ATCGACGTCAAGAAAGTTCCTCCATCTCCATTGGCATCATCGTGTTCAGCTACACGTCTCAG 1071  
QY 961 ATCTTCTGCTTTCGTGGAGGCAATATACAGAGCCAGCGAGTTCCTGTCATGATGATG 1020  
DB 1072 ATCTTCTGCTTTCGTGGAGGCAATATACAGAGCCAGCGAGTTCCTGTCATGATGATG 1131  
QY 1021 AACTGGAGCAGATCCGACGCTCGTCTCAAGGGCTCTTTCGGCTCGCTGCTGCTACCTC 1080  
DB 1132 AACTGGAGCAGATCCGACGCTCGTCTCAAGGGCTCTTTCGGCTCGCTGCTGCTACCTC 1191  
QY 1081 ACCTGGGCCACAGACCAAGAGAGTTCATCAAGATAACCTGCGCGCTCCATCCGCGCC 1140  
DB 1192 ACCTGGGCCACAGACCAAGAGAGTTCATCAAGATAACCTGCGCGCTCCATCCGCGCC 1251  
QY 1141 GTGGTCAACATCTTTCGTGGGCAAGGCGCTGTGTCTTATCTCTGCGCATTTCTTTGGC 1200  
DB 1252 GTGGTCAACATCTTTCGTGGGCAAGGCGCTGTGTCTTATCTCTGCGCATTTCTTTGGC 1311  
QY 1201 GCTGTGAGGTGTGGAGAGTGTCTTCCAGAGGAGGAGCGCGCTTTTCCCGGCC 1260  
DB 1312 GCTGTGAGGTGTGGAGAGTGTCTTCCAGAGGAGGAGCGCGCTTTTCCCGGCC 1371  
QY 1261 TGCTACAGCGCGACGCGCGCTGAAAGTCTCTGGGGCTGACGCTGCGCTGCGGCTCGTC 1320  
DB 1372 TGCTACAGCGCGACGCGCGCTGAAAGTCTCTGGGGCTGACGCTGCGCTGCGGCTCGTC 1431  
QY 1321 GTCTTCAAGCTGTCTATGCGCATTTATGTGCGCATTTTCGCGTGTCTATGCGGCTCACC 1380  
DB 1432 GTCTTCAAGCTGTCTATGCGCATTTATGTGCGCATTTTCGCGTGTCTATGCGGCTCACC 1491  
QY 1381 GCGAGCTCACGCGCGCGCGCTCTGTCTTCTGCTGCCAGCTCTTTCACCTGCGGCTG 1440  
DB 1492 GCGAGCTCACGCGCGCGCGCTCTGTCTTCTGCTGCCAGCTCTTTCACCTGCGGCTG 1551  
QY 1441 CTCTGCGCAAGCTGTGTGGCACCAAGTCTTCTTTCGAGCTGCGCATTTTCGTCATCGGC 1500  
DB 1552 CTCTGCGCAAGCTGTGTGGCACCAAGTCTTCTTTCGAGCTGCGCATTTTCGTCATCGGC 1611  
QY 1501 GGATCTGACAGCTGTCCGGCTTGTGTGCACTCCCTCGAGGGCTCTATTCGAAGCTTACCGA 1560  
DB 1612 GGATCTGACAGCTGTCCGGCTTGTGTGCACTCCCTCGAGGGCTCTATTCGAAGCTTACCGA 1671  
QY 1561 ACCAAGCGGAGGACTAG 1578  
DB 1672 ACCAAGCGGAGGACTAG 1689

RESULT 12  
AX497266

source	1. .1734	
	/organism="Homo sapiens"	
	/mol_type="unassigned DNA"	
	/db_xref="taxon:9606"	
CDS	112..1689	
	/note="unnamed protein product"	
	/codon_start=1	
	/protein_id="CAD48687.1"	
	/db_xref="GI:23342646"	
	/translation="NATLLRSKLSNVATSVSNKSQAKSMGFARMFQAAATDEEAVGF	
	AHCDDLEFHRGLQMDILKABGEPGDEGAEAPVEGDIHYORGSGAPLPSPGSKDQV	
	GGGFGFHGDKPKITAWAEGWNVTAIQGMVLGYPYAILHGGYLGLELLIPAAVCC	
	YTKLILACLIEENEDSEVRVDSVAIANACCAPRPTLGRVNVNAQIIEIIVMTC	
	ILYVVNSGLMYSNFPGLPVSKQSIITATVLLPCLKLNKAVSKFSLCTLAHFV	
	INILVAVCLSPARDWAEKVFIDVKKFPIISIGIIVFSYTSQIFLPSLEGNMOOPS	
	EPhCMNWHIAACVLKGLPALVALVLTWADETKEAITDNLPGSIRAVVNIFLVAKALL	
	SYPLPFFAAVEVLEKSLFQEGSRAPFPACYSGDRKLSWGLRLCALVFFILLMAIYV	
	PFPALLMGLTSLTGAGLCFLPLPSLFIHLRLMLWRKLLMHQVFDVAIFVIGGCVSGF	
	VHSLEGLIEAYRTNAED"	
ORIGIN		
Query Match	99.8%; Score 1574.8; DB 6; Length 1734;	
Best Local Similarity	99.9%; Pred. No. 4.7e-229;	
Matches 1576; Conservative	0; Mismatches 2; Indels 0; Gaps 0;	
QY	1	ATGGCCACCTTGTCTCGCAGCAAGCTGTCCAAGTGGCCACGTCGCTGTCCAAAGTCC 60
DB	112	ATGGCCACCTTGTCTCGCAGCAAGCTGTCCAAGTGGCCACGTCGCTGTCCAAAGTCC 171
QY	61	CAGGCCAAGATGAGCGGCAATGTCGCCAGGATGGGTTTTTCAGCGGCGCCACGATGAGGAG 120
DB	172	CAGGCCAAGATGAGCGGCAATGTCGCCAGGATGGGTTTTTCAGCGGCGCCACGATGAGGAG 231
QY	121	GGGTGGGCTTCGCGCAATTCGACGACCTCGACTTTTGAGACCGCGAGGCGCTGCAGATG 180
DB	232	GGGTGGGCTTCGCGCAATTCGACGACCTCGACTTTTGAGACCGCGAGGCGCTGCAGATG 291
QY	181	GACATCTCAAGACCGAGGAGAGCCCTCGCGGAGCGAGGCGCTGAAGCGCCGCTCGAG 240
DB	292	GACATCTCAAGACCGAGGAGAGCCCTCGCGGAGCGAGGCGCTGAAGCGCCGCTCGAG 351
QY	241	GGAGACATCCATTTATCAGCGAGGACGAGCGAGCTCTCTGCGCGCTTCGCGCTCCAAAGGAC 300
DB	352	GGAGACATCCATTTATCAGCGAGGACGAGCGAGCTCTCTGCGCGCTTCGCGCTCCAAAGGAC 411
QY	301	CAGGTGGGAGTGGTGGCGAATTCGGGGGCGACGACCAAGCCCAAAATACCGCGGTGGAG 360
DB	412	CAGGTGGGAGTGGTGGCGAATTCGGGGGCGACGACCAAGCCCAAAATACCGCGGTGGAG 471
QY	361	GCAGGCTGGAACTGACCAACGCTATCCAGGCGATGTCGTGTGGGCGCTACCCCTACGCGC 420
DB	472	GCAGGCTGGAACTGACCAACGCTATCCAGGCGATGTCGTGTGGGCGCTACCCCTACGCGC 531
QY	421	ATCCTGACGCGCGGCTTACCTGGGGTGTGTTTCTCATCATCTTCGCGCGCTTGTGTGCTGC 480
DB	532	ATCCTGACGCGCGGCTTACCTGGGGTGTGTTTCTCATCATCTTCGCGCGCTTGTGTGCTGC 591
QY	481	TACACGGCAAGATCCTCATCGCGTGCCTGTACGAGGAGATGAAGACGCGGAGGTGGTG 540
DB	592	TACACGGCAAGATCCTCATCGCGTGCCTGTACGAGGAGATGAAGACGCGGAGGTGGTG 651
QY	541	CGGTCGGGAGCTCGTACGTGGGCGCATAGCCACGCTGCTGCGCGCGCTTCCCAACG 600
DB	652	CGGTCGGGAGCTCGTACGTGGGCGCATAGCCACGCTGCTGCGCGCGCTTCCCAACG 711
QY	601	CTGGGCGGCGAGTGGTGAACCTAGCGCAGATCATCGAGCTGGTGTGATGACGTGCACTCTG 660
DB	712	CTGGGCGGCGAGTGGTGAACCTAGCGCAGATCATCGAGCTGGTGTGATGACGTGCACTCTG 771
QY	661	TACGTGGTGTGATGGCAACCTCATGTACACAGCTTCCCGGGGCTGCGCGTGTGCGAG 720
DB	772	TACGTGGTGTGATGGCAACCTCATGTACACAGCTTCCCGGGGCTGCGCGTGTGCGAG 831
QY	721	AAGTCTCTGGTCCATTTATCGCCACGCGCGTGTGCTGCTTGGCCTTCTTAAAGAACCTC 780
DB	832	AAGTCTCTGGTCCATTTATCGCCACGCGCGTGTGCTGCTTGGCCTTCTTAAAGAACCTC 891
QY	781	AAGGCGGTGTCAAGTTCAAGTCTGCTGTGCACTCTGGGCCACTTCTGTCATCAATATCTTG 840
DB	892	AAGGCGGTGTCAAGTTCAAGTCTGCTGTGCACTCTGGGCCACTTCTGTCATCAATATCTTG 951
QY	841	GTCATAGCCTACTGTCTATTCGCGGCGCGCGACTGGGGCTCTGGAGAAAGGTCAAGTTCTTAC 900
DB	952	GTCATAGCCTACTGTCTATTCGCGGCGCGCGACTGGGGCTCTGGAGAAAGGTCAAGTTCTTAC 1011
QY	901	ATCGACGTCAAGAAGTTCCCATCTCCATTTGGCATCATCTGTTTTCAGTACACGCTCTCAG 960
DB	1012	ATCGACGTCAAGAAGTTCCCATCTCCATTTGGCATCATCTGTTTTCAGTACACGCTCTCAG 1071
QY	961	ATCTTCTCTGCTTCTGCTGAGGGCAATATGACGAGCCGACGAGTTTCCACTGCAATGATG 1020
DB	1072	ATCTTCTCTGCTTCTGCTGAGGGCAATATGACGAGCCGACGAGTTTCCACTGCAATGATG 1131
QY	1021	AACTGAGCGCACATCGCAGCGCTGCTCAAGGGCTCTTTCGCGCTCGTTCGCTTACCTC 1080
DB	1132	AACTGAGCGCACATCGCAGCGCTGCTCAAGGGCTCTTTCGCGCTCGTTCGCTTACCTC 1191
QY	1081	ACCTGGGCGCGACGACCAAGAGGTATCATCGGATAAAGTCTGCGGCTTCCATCCGCGCC 1140
DB	1192	ACCTGGGCGCGACGACCAAGAGGTATCATCGGATAAAGTCTGCGGCTTCCATCCGCGCC 1251
QY	1141	GTGTGCAACATCTTCTGCTGGCCAAAGGCTGTTGCTATCTTCTGCTTCTTTCGCTTTCG 1200
DB	1252	GTGTGCAACATCTTCTGCTGGCCAAAGGCTGTTGCTATCTTCTGCTTCTTTCGCTTTCG 1311
QY	1201	GCTGTCGAGGTGCTGAGGAAGTCTCTTTCAGGAAGGAGCGCGGCTTTCCTTTCGCGGC 1260
DB	1312	GCTGTCGAGGTGCTGAGGAAGTCTCTTTCAGGAAGGAGCGCGGCTTTCCTTTCGCGGC 1371
QY	1261	TGCTACAGCGGCGACGCGGCGCTGAAAGTCTTGGGGGCTGACGCTGCGCTGCGCGCTCGTC 1320
DB	1372	TGCTACAGCGGCGACGCGGCGCTGAAAGTCTTGGGGGCTGACGCTGCGCTGCGCGCTCGTC 1431
QY	1321	GTCCTACGCTGCTCATGGCCATTTATGTCGCGCACTTTCGCGCTGCTCATGGGCTTCC 1380
DB	1432	GTCCTACGCTGCTCATGGCCATTTATGTCGCGCACTTTCGCGCTGCTCATGGGCTTCC 1491
QY	1381	GGCAGCTCAGCGGCGCGGCTCTGTTTCTTGTGTCGCCAGGCTTTCCTTTCACCTGCGGCTG 1440
DB	1492	GGCAGCTCAGCGGCGCGGCTCTGTTTCTTGTGTCGCCAGGCTTTCCTTTCACCTGCGGCTG 1551
QY	1441	CTTCTGGCGCAAGCTGCTGTGGCACCAAGTCTTCTTTCGACGTGCGCATCTTTCGTCATCGGC 1500
DB	1552	CTTCTGGCGCAAGCTGCTGTGGCACCAAGTCTTCTTTCGACGTGCGCATCTTTCGTCATCGGC 1611
QY	1501	GGCATCTGAGGCTGTCGGGCTTCTGCACTCCCTCGAGGGGCTCATCGAAGCTTACCGA 1560
DB	1612	GGCATCTGAGGCTGTCGGGCTTCTGCACTCCCTCGAGGGGCTCATCGAAGCTTACCGA 1671
QY	1561	ACCAACGCGGAGGACTAG 1578
DB	1672	ACCAACGCGGAGGACTAG 1689
RESULT 11		
AX497272		
LOCUS	Sequence 39 from Patent WO0229058.	
DEFINITION	AX497272	
ACCESSION	AX497272.1	
VERSION	GI:23342647	
KEYWORDS	Homo sapiens (human)	
SOURCE	Homo sapiens	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE	1	
	PAT 26-SEP-2002	

AHCDLDLFEHQQLQMDILKAEGPCDEGAEPVEGDHYQSGAPLPPSSKQDV  
GGGEGFGHDKPKITAEAGWVNTNATQGMFVLGPVAILHGGYGLFLFIIFAAVTC  
YTKLILIAELYENEDGVRDVSAAIANACAPREFPTLGGRVNVAOIIELVNVC  
ILVYVSGNLMYNSFPCLPVSKSMISIIATAVLLPCAFLEKXAKVSKESLLCTIAFPV  
INILVATCLSRARDNAWEKVIYDVKFPPISIGIIVFYSTQIFDPSLEGNNQOPS  
EFHGMNTHIAACVLKGLPALVAYLUTWDETBKAITDNLPGSIRAVNIFLVAKALL  
SYPLPFAAVTEVLEKSLFQBSRAFFPACYSGDRLKSLGLTLRICALVWFTLLMAYV  
PHFALLMGLTGSITGAGLCFLPLSLFHLRLWKLWHQVFDVAIFVIGIGCSVSGF  
VHSLEGLIEAYRTNAED"

ORIGIN

Query Match 99.8%; Score 1574.8; DB 6; Length 1734;  
Best Local Similarity 99.9%; Pred. No. 4.7e-229;  
Matches 1576; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	1	ATGGCCACCTTGCTCCGACGACAGCTGTCCAAAGTGGCCACGCTCCGTGTCCAAAGTCC	60
Db	112	ATGGCCACCTTGCTCCGACGACAGCTGTCCAAAGTGGCCACGCTCCGTGTCCAAAGTCC	171
QY	61	CAGGCCAAGATGAGCGGCATGTTCCGACGATGGGTTTTCAGGCGGCCACGGATGAGGAG	120
Db	172	CAGGCCAAGATGAGCGGCATGTTCCGACGATGGGTTTTCAGGCGGCCACGGATGAGGAG	231
QY	121	GCGGTGGCTTCGCGCATTCGACGACCTCGACTTTGAGCACCGCCAGGCGCTTCGACATG	180
Db	232	GCGGTGGCTTCGCGCATTCGACGACCTCGACTTTGAGCACCGCCAGGCGCTTCGACATG	291
QY	181	GACATCTGAAGCCGAGGAGAGCCCTCGCGGAGCGAGGCGCTGAAGCGCCGCTCGAG	240
Db	292	GACATCTGAAGCCGAGGAGAGCCCTCGCGGAGCGAGGCGCTGAAGCGCCGCTCGAG	351
QY	241	GGAGACATCATTTATCAGCGAGGAGCGGAGCTCTCTCGCGCCCTCCGGCTCCAAAGGAC	300
Db	352	GGAGACATCATTTATCAGCGAGGAGCGGAGCTCTCTCGCGCCCTCCGGCTCCAAAGGAC	411
QY	301	CAGGTGGAGGTGGTGGCGAATTCGGGGGCCACGACAGCCCAAAATCACGGGCTGGGAG	360
Db	412	CAGGTGGAGGTGGTGGCGAATTCGGGGGCCACGACAGCCCAAAATCACGGGCTGGGAG	471
QY	361	GCAGGTGGAAAGTACCAACGCGATCCAGGCGATGTTTGTGTGGGCTTACCTACGCC	420
Db	472	GCAGGTGGAAAGTACCAACGCGATCCAGGCGATGTTTGTGTGGGCTTACCTACGCC	531
QY	421	ATCTGCACGCGCGCTACCTGGGGTTGTTTCTCATCATCTTCGCGCGCTGTGTGCTGC	480
Db	532	ATCTGCACGCGCGCTACCTGGGGTTGTTTCTCATCATCTTCGCGCGCTGTGTGCTGC	591
QY	481	TACACCGGCAAGATCTCTCATCGCGTGCCTGTACGAGGAGAAATGAAGCGCGAGGTGGTG	540
Db	592	TACACCGGCAAGATCTCTCATCGCGTGCCTGTACGAGGAGAAATGAAGCGCGAGGTGGTG	651
QY	541	CGCGTGGGAGCTGTACGTGGCGCATAGCAACGCTGTGCGCGCGCGCTTCCCAACG	600
Db	652	CGCGTGGGAGCTGTACGTGGCGCATAGCAACGCTGTGCGCGCGCGCTTCCCAACG	711
QY	601	CTGCGGCGCGAGTGTGTGACGTAGCGAGATCATCGAGCTGTGTGATGATGATGATGATG	660
Db	712	CTGCGGCGCGAGTGTGTGACGTAGCGAGATCATCGAGCTGTGTGATGATGATGATGATG	771
QY	661	TACGTGTGTGTGAGTGGCAACCTCATGTACAAAGCTTCCCGGGGCTGCGCGCTGTGCGAG	720
Db	772	TACGTGTGTGTGAGTGGCAACCTCATGTACAAAGCTTCCCGGGGCTGCGCGCTGTGCGAG	831
QY	721	AAGTCTGTGTCATTAATTCGCCACGCGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	780
Db	832	AAGTCTGTGTCATTAATTCGCCACGCGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	891
QY	781	AAGCGGTGTGCCAAGTTTCAGTCTGTGTGACCTGTGCGCCACCTTCGTATCATATATCTG	840
Db	892	AAGCGGTGTGCCAAGTTTCAGTCTGTGTGACCTGTGCGCCACCTTCGTATCATATATCTG	951
QY	841	GTCCATAGCTTACTGTCTATCGCGGCGCGACTGGGCGCTGGGAGAGGTCAAGTTCTTAC	900

Db	952	GTCCATAGCTTACTGTCTATCGCGGGCGCGGACTGGGCGCTGGGAGAGGTCAAGTTCTTAC	1011
QY	901	ATCGAGCTCAAGAAGTTCCCATCTCCATTTGGCATCATCGTGTTCAGCTACAGTCTCAG	960
Db	1012	ATCGAGCTCAAGAAGTTCCCATCTCCATTTGGCATCATCGTGTTCAGCTACAGTCTCAG	1071
QY	961	ATCTTCTCGCTTTCGCTGGAGGGCAATATGCACGACGCCACCGAGTTCACATGCATGATG	1020
Db	1072	ATCTTCTCGCTTTCGCTGGAGGGCAATATGCACGACGCCACCGAGTTCACATGCATGATG	1131
QY	1021	AACTGACGACATCGAGCTCGCTCAAGGGCTCTTTCGCGCTGTGCGCTACCTC	1080
Db	1132	AACTGACGACATCGAGCTCGCTCAAGGGCTCTTTCGCGCTGTGCGCTACCTC	1191
QY	1081	ACCTGGGCGACGAGACCAAGAGGTATACACGATAACTGCGCGGCTCCATCCGCGCC	1140
Db	1192	ACCTGGGCGACGAGACCAAGAGGTATACACGATAACTGCGCGGCTCCATCCGCGCC	1251
QY	1141	GTGGTCAACATCTTCTGGTGGCAAGGGCTGTGTCTTCTATCTCTGCGCATTTCTTGGC	1200
Db	1252	GTGGTCAACATCTTCTGGTGGCAAGGGCTGTGTCTTCTATCTCTGCGCATTTCTTGGC	1311
QY	1201	GCTGTGAGGTGCTGAGAAAGTCTTCCAGAAAGGAGCGCGCTTTTTCGCGGCC	1260
Db	1312	GCTGTGAGGTGCTGAGAAAGTCTTCCAGAAAGGAGCGCGCTTTTTCGCGGCC	1371
QY	1261	TGCTACAGCGGCGACGGGCGCTTGAAGTCTGGGGGCTGACGCTGCGCTGCGCGCTGCTC	1320
Db	1372	TGCTACAGCGGCGACGGGCGCTTGAAGTCTGGGGGCTGACGCTGCGCTGCGCGCTGCTC	1431
QY	1321	GTCTTACGCTGCTCATGCGCATTTATGTGCCGCACTTTCGCGCTGTCTATGGGCTTCAAC	1380
Db	1432	GTCTTACGCTGCTCATGCGCATTTATGTGCCGCACTTTCGCGCTGTCTATGGGCTTCAAC	1491
QY	1381	GGCAGCTCAAGCGCGCGCTTCTGTTTCTGCTGCCAGGCTTTTCACTGCGCGCTG	1440
Db	1492	GGCAGCTCAAGCGCGCGCTTCTGTTTCTGCTGCCAGGCTTTTCACTGCGCGCTG	1551
QY	1441	CTTGTGGCGAGCTGTGTGGCAACCAAGTCTTTCGAGCTGCGCATCTTGTCTATCGCGC	1500
Db	1552	CTTGTGGCGAGCTGTGTGGCAACCAAGTCTTTCGAGCTGCGCATCTTGTCTATCGCGC	1611
QY	1501	GGCATCTGACGCTGCTGCGCTTTCGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1560
Db	1612	GGCATCTGACGCTGCTGCGCTTTCGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1671
QY	1561	ACCAACGCGGAGGACTAG 1578	
Db	1672	ACCAACGCGGAGGACTAG 1689	

RESULT 10

AX497270	1734 bp	DNA	linear	PAT 26-SEP-2002
Sequence 37 from Patent WO0229058.				
AX497270				
AX497270.1	GI:23342645			
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
Homo sapiens (human)				
Homo sapiens				
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
1				
Shinketsu, R.A., Taupier, R.J., Burgess, C.E., Zerhusen, B.D.,				
Mezes, P.S., Rastelli, L., Malyankar, U.M., Grosse, W.M.,				
Alsobrook, J.P., Lepley, D.M., Spytek, K.A., Li, L., Edinger, S.,				
Gerlach, V., Ellerman, K., Macdougall, J., Gunther, E., Millet, I.,				
Stone, D., Ellerman, K., Macdougall, J., Gunther, E., Millet, I.,				
Human proteins, polynucleotides encoding them and methods of using				
the same				
Patent: WO 0229058-A 37 11-APR-2002;				
Curagen Corporation (US)				
Location/Qualifiers				
FEATURES				

Query Match 99.8%; Score 1574.8; DB 6; Length 1734;  
Best Local Similarity 99.9%; Pred. No. 4.7e-229;  
Matches 1576; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCACCTTCTCGCAGCAAGCTGTCACAGTGGCCACCTCGTGTCCAAACAGTCC 60  
DB 112 ATGGCCACCTTCTCGCAGCAAGCTGTCACAGTGGCCACCTCGTGTCCAAACAGTCC 171

QY 61 CAGGCCAGATGAGCGGCATGTTCCAGGATGGTTTTCAGCGCGCCACGATGAGGAG 120  
DB 172 CAGGCCAGATGAGCGGCATGTTCCAGGATGGTTTTCAGCGCGCCACGATGAGGAG 231

QY 121 CGCGTGGGCTTCGCGCATTTGACACGACCTTCGATTTTGAACCGCAGCGGCTTCAGATG 180  
DB 232 CGCGTGGGCTTCGCGCATTTGACACGACCTTCGATTTTGAACCGCAGCGGCTTCAGATG 291

QY 181 GACATCTGAAGCCGAGGAGAGCCCTCGCGGAGAGGGCGCTGGAAGCGCCGCTCGAG 240  
DB 292 GACATCTGAAGCCGAGGAGAGCCCTCGCGGAGAGGGCGCTGGAAGCGCCGCTCGAG 351

QY 241 GGAGACATCATATCAGCGGAGCGAGCTCTCTCGCGCCCTCGGCTCCAGGAC 300  
DB 352 GGAGACATCATATCAGCGGAGCGAGCTCTCTCGCGCCCTCGGCTCCAGGAC 411

QY 301 CAGGTGGGAGGTGGTGGCGAATTCGGGGGCCACGACAGCCCAAAATCACGGCGTGGGAG 360  
DB 412 CAGGTGGGAGGTGGTGGCGAATTCGGGGGCCACGACAGCCCAAAATCACGGCGTGGGAG 471

QY 361 CGAGGCTGGAACGTGACCAACGCCATCCAGGGCATGTTCTGCTCGGCGCTACCTACGCC 420  
DB 472 CGAGGCTGGAACGTGACCAACGCCATCCAGGGCATGTTCTGCTCGGCGCTACCTACGCC 531

QY 421 ATCTGTCAGCGGCGCTACCTGGGTTGTTTCTCATCATCTTCGCGCGCTGTGTGCTGC 480  
DB 532 ATCTGTCAGCGGCGCTACCTGGGTTGTTTCTCATCATCTTCGCGCGCTGTGTGCTGC 591

QY 481 TACACCGCAAGATCTCATCGCTGCTGTAGAGGAGAAATGAAGACGGCGAGGTGGTG 540  
DB 592 TACACCGCAAGATCTCATCGCTGCTGTAGAGGAGAAATGAAGACGGCGAGGTGGTG 651

QY 541 CGCGTGGGAGCTCGTACGTGGCCATAGCAACGCTGCTCGCGCCCGCGCTTCCCAACG 600  
DB 652 CGCGTGGGAGCTCGTACGTGGCCATAGCAACGCTGCTCGCGCCCGCGCTTCCCAACG 711

QY 601 CTGGGCGCGGAGTGTGAAGCTAGCGGAGATCATCGAGCTGCTGATGACGTGCATCCTG 660  
DB 712 CTGGGCGCGGAGTGTGAAGCTAGCGGAGATCATCGAGCTGCTGATGACGTGCATCCTG 771

QY 661 TAGCTGCTGTAGTGGCAACCTCATGTACACAGCTTCCCGGGCTGCGCGTGTGCGAG 720  
DB 772 TAGCTGCTGTAGTGGCAACCTCATGTACACAGCTTCCCGGGCTGCGCGTGTGCGAG 831

QY 721 AAGTCTGTGTCATTATCCGACGCGCGTGTGCTGCTGCTGCGCTTCCCTTAAGAACCTC 780  
DB 832 AAGTCTGTGTCATTATCCGACGCGCGTGTGCTGCTGCTGCGCTTCCCTTAAGAACCTC 891

QY 781 AAGCGCGTGTCAAGTTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840  
DB 892 AAGCGCGTGTCAAGTTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 951

QY 841 GTCATAGCTTACGTCTATCCGCGGCGCGAGCTGCGGCTGGGAGAGGTCAAGTTCTAC 900  
DB 952 GTCATAGCTTACGTCTATCCGCGGCGCGAGCTGCGGCTGGGAGAGGTCAAGTTCTAC 1011

QY 901 ATCGAGCTCAAGAGTTCCCATCTCCATTTGGCATCATGCTGTTTTCAGCTACAGCTCTCAG 960  
DB 1012 ATCGAGCTCAAGAGTTCCCATCTCCATTTGGCATCATGCTGTTTTCAGCTACAGCTCTCAG 1071

QY 961 ATCTTCTGCTTTCGCTGGAGGGCAATATGACAGCCAGCCAGCGAGTTCACCTGATGATG 1020  
DB 1072 ATCTTCTGCTTTCGCTGGAGGGCAATATGACAGCCAGCCAGCGAGTTCACCTGATGATG 1131

RESULT 9  
AX497268  
LOCUS AX497268 1734 bp DNA linear PAT 26-SEP-2002  
DEFINITION Sequence 35 from Patent WO0229058.  
ACCESSION AX497268  
VERSION AX497268.1 GI:23342643  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Shinkets, R.A., Taupier, R.J., Burgess, C.E., Zerhusen, B.D.,  
Mezes, P.S., Rastelli, L., Malyankar, U.M., Grosse, W.M.,  
Alsobrook, J.P., Lepley, D.M., Spytek, K.A., Li, L., Edinger, S.,  
Gerlach, V., Billeman, K., Macdougall, J., Gunther, E., Millet, I.,  
Stone, B., Smithson, G. and Szekeres, E.S.  
TITLE Human proteins, polynucleotides encoding them and methods of using  
the same  
JOURNAL Patent: WO 0229058-A 35 11-APR-2002;  
Curagen Corporation (US)  
FEATURES  
source  
1..1734  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
112..1689  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD48686.1"  
/db\_xref="GI:23342644"  
/translation="MATLLRSKLSNVATSVSNKSQAKMSGFARMFQAATDEAVGF"

Db 160 CAGGCCAAGATGAGCGGCACTGTTCCGAGGATGGGTTTTCAGGCGGCCACGGATGAGGAG 219  
QY 121 GCGGTGGGCTTCGGGCACTTCGACGACTCGACTTTTCAGCACCGCCAGGCGCTTCGACATG 180  
Db 220 GCGGTGGGCTTCGGGCACTTCGACGACTCGACTTTTCAGCACCGCCAGGCGCTTCGACATG 279  
QY 181 GACATCTCTGAAGCCGAGGAGAGCCCTCGGGGACGAGGCGCTGAAGCGCCGCTCGAG 240  
Db 280 GACATCTGAAGCCGAGGAGAGCCCTCGGGGACGAGGCGCTGAAGCGCCGCTCGAG 339  
QY 241 GGAGACATCATATTCAGCGAGGAGCGGAGCTCTTCGCGGCTTCGCGCTTCAAGGAC 300  
Db 340 GGAGACATCATATTCAGCGAGGAGCGGAGCTCTTCGCGGCTTCGCGCTTCAAGGAC 399  
QY 301 CAGGTGGAGGTGGTGGCAATTCGGGGCCACGACCAAGCCCAAAATCACGGGCTGGAG 360  
Db 400 CAGGTGGAGGTGGTGGCAATTCGGGGCCACGACCAAGCCCAAAATCACGGGCTGGAG 459  
QY 361 GCAGGCTGGAACTGACCAACGCGCATCCAGGGCATGTTCTGCTGGGCTTACCTACGCC 420  
Db 460 GCAGGCTGGAACTGACCAACGCGCATCCAGGGCATGTTCTGCTGGGCTTACCTACGCC 519  
QY 421 ATCTGTCAACGGGCTTACCTGGGCTTGTCTTCATCATCTTCGCGCGCTTGTGTCTGC 480  
Db 520 ATCTGTCAACGGGCTTACCTGGGCTTGTCTTCATCATCTTCGCGCGCTTGTGTCTGC 579  
QY 481 TACACGGCAAGATCCTCATCGGCTGCTGTACGAGGAGATGAAGACGGGAGTGGTG 540  
Db 580 TACACGGCAAGATCCTCATCGGCTGCTGTACGAGGAGATGAAGACGGGAGTGGTG 639  
QY 541 CGGCTGGGACTCGTACGTGGCCATAGCAACGCTGCTGCGCGCCCGGCTTCCCAACG 600  
Db 640 CGGCTGGGACTCGTACGTGGCCATAGCAACGCTGCTGCGCGCCCGGCTTCCCAACG 699  
QY 601 CTGGGCGGCGAGTGGTGAACGCTAGCGAGATCATCGAGTGGTGAACGCTAGCGAGTGGT 660  
Db 700 CTGGGCGGCGAGTGGTGAACGCTAGCGAGATCATCGAGTGGTGAACGCTAGCGAGTGGT 759  
QY 661 TAGT 720  
Db 760 TAGT 819  
QY 721 AAGTCTGT 780  
Db 820 AAGTCTGT 879  
QY 781 AAGGCGGT 840  
Db 880 AAGGCGGT 939  
QY 841 GTCATAGCTTACTGTCTATCGGGGCGCGCATGCGGCTGGGAGAGTCAAGTGTCTAC 900  
Db 940 GTCATAGCTTACTGTCTATCGGGGCGCGCATGCGGCTGGGAGAGTCAAGTGTCTAC 999  
QY 901 ATGAGCTCAAGAGTTCCCATCTCCATTTGGGATCATCGTGTTCAGCTACAGCTTCTAC 960  
Db 1000 ATGAGCTCAAGAGTTCCCATCTCCATTTGGGATCATCGTGTTCAGCTACAGCTTCTAC 1059  
QY 961 ATCTTCTGCTTCTGCTGTGAGGCAATATGACGAGCCGACGAGTTCACATCATGATG 1020  
Db 1060 ATCTTCTGCTTCTGCTGTGAGGCAATATGACGAGCCGACGAGTTCACATCATGATG 1119  
QY 1021 AACTGAGCAGCATTCGAGCTCGCTGTCTCAAGGCGCTTTCGCGCTTCGCTACCTC 1080  
Db 1120 AACTGAGCAGCATTCGAGCTCGCTGTCTCAAGGCGCTTTCGCGCTTCGCTACCTC 1179  
QY 1081 ACCTGGGCGCAGACGAGGAGTTCATCAGGATAACTGCGCGGCTTCCATCCGCGC 1140  
Db 1180 ACCTGGGCGCAGACGAGGAGTTCATCAGGATAACTGCGCGGCTTCCATCCGCGC 1239  
QY 1141 GTGGTCAACATCTTTCTGTGGTGGCCAGGCGCTGTGTCTTATCTCTGCTTCTTTGCC 1200

1240 GTGGTCAACATCTTTCTGTGGCCAAAGGCGCTGTGTCTTATCTCTGCGCATCTTTTGCC 1299  
QY 1201 GCTGTGAGGTGCTGTGAGAAAGTCTCTTCCAGAAAGGACGCGCGCTTTTCCCGGCC 1260  
Db 1300 GCTGTGAGGTGCTGTGAGAAAGTCTCTTCCAGAAAGGACGCGCGCTTTTCCCGGCC 1359  
QY 1261 TGCTACAGCGGCGACGCGGCGCTTGAAGTCTCTGGGGGTGACGCTGCGCTCGCGCTGCTC 1320  
Db 1360 TGCTACAGCGGCGACGCGGCGCTTGAAGTCTCTGGGGGTGACGCTGCGCTCGCGCTGCTC 1419  
QY 1321 GTCTTACGCTGCTCATGGCCATTTATGTCCGCACTTTCGCGCTGCTCATGGGCTCAC 1380  
Db 1420 GTCTTACGCTGCTCATGGCCATTTATGTCCGCACTTTCGCGCTGCTCATGGGCTCAC 1479  
QY 1381 GGCAGGCTCACGCGGCGCGGCGCTTGTCTTCTGCTGCCAGCCTTCTTACCTGCGCGCTG 1440  
Db 1480 GGCAGGCTCACGCGGCGCGGCGCTTGTCTTCTGCTGCCAGCCTTCTTACCTGCGCGCTG 1539  
QY 1441 CTCTGGCGCAAGCTGCTGTGGCACCAAGTCTTCTTGCAGCTCGCCATCTTTCGCTATCGGC 1500  
Db 1540 CTCTGGCGCAAGCTGCTGTGGCACCAAGTCTTCTTGCAGCTCGCCATCTTTCGCTATCGGC 1599  
QY 1501 GGCATCTGACGCTGCTCGGCTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560  
Db 1600 GGCATCTGACGCTGCTCGGCTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1659  
QY 1561 ACCAAGCGGAGGACTAG 1578  
Db 1660 ACCAAGCGGAGGACTAG 1677

RESULT 8  
AX497264  
LOCUS  
DEFINITION  
Sequence 31 from Patent WO0229058.  
ACCESSION  
AX497264  
VERSION  
AX497264.1 GI:23342640  
KEYWORDS  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
1  
AUTHORS  
Shinkets, R.A., Taupier, R.J., Burgees, C.E., Zerhusen, B.D.,  
Mezes, P.S., Rastelli, L., Malyankar, U.M., Grosse, W.M.,  
Alsobrook, J.P., Lepley, D.M., Spytek, K.A., Li, L., Edinger, S.,  
Gerlach, V., Ellerman, K., Macdougall, J., Gunther, E., Millet, I.,  
Stone, D., Smithson, G. and Szekeres, E.S.  
Human proteins, polynucleotides encoding them and methods of using  
the same  
Patent: WO 0229058-A 31 11-APR-2002;  
Curagen Corporation (US)

TITLE  
Location/Qualifiers  
1. .1734  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
112. .1689  
/note="unnamed protein product"

JOURNAL  
Patent: WO 0229058-A 31 11-APR-2002;  
Curagen Corporation (US)

FEATURES  
source  
1. .1734  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
112. .1689  
/note="unnamed protein product"

CDS  
112. .1689  
/codon\_start=1  
/protein\_id="CAD48685.1"  
/db\_xref="GI:23342641"

ORIGIN  
1  
translation="MATLRSKLSNVATSVSNKSQAKMSGMFARMGFOATDEEAVGF  
AHCDDLDFHRQLQMDILKAEFGPCDEGAPVEGDHYHQSGAPLPPSGSKDQV  
GGGFGHDHFKPTAWAEGWNVTAIOGMFVLGYPYAILHGRVNVVAQIIELAVTCC  
YTGKILIACLYEENEDGEVVRVDSYVIANACCAFPPTLGGRVNVVAQIIELAVTCC  
ILYVVSNGLMVNSPGLPVSKWSIIATAVLLPCALFKLKAWSKSLCTLAHPV  
INILVAYCLSRARDWAEKVPYIDVKFPISIGIIVFSYISOIFLPLEGNMQOPS  
EHEMNWTHIAACVLKGLFALVYLTWADETKEATDNLPGSIRAVNVIPVMAKILY  
SYPLPFAAVEVLEKSLFOGSRAPFPCYSGDGRKLSWGLTRLCRALVNVFTLMAIYV  
PHFALIMGLTSLTGAGLCFLPLSLFHLRLWRKLMHQQVFDVAIFVIGGICSVSGF  
VHSLGLEIYARTNAED"

Db 261 GCGGTGGGCTTCGCGATTGCGAGCACTCGACTTTGAGCACCGCGCTCGAGATG 320  
Qy 181 GACATCTTGAAGCGGAGAGAGCCCTGCGGGACGAGGGCGCTGAAGCCCGCTCGAG 240  
Db 321 GACATCTTGAAGCGGAGAGAGCCCTGCGGGACGAGGGCGCTGAAGCCCGCTCGAG 380  
Qy 241 GAGACATCTTGAAGCGGAGAGAGCCCTGCGGGACGAGGGCGCTGAAGCCCGCTCGAG 300  
Db 381 GAGACATCTTGAAGCGGAGAGAGCCCTGCGGGACGAGGGCGCTGAAGCCCGCTCGAG 440  
Qy 301 GAGGTGGAGGTGGTGGGAAATTCGGGGGCGACGACAAAGCCCAAAATCAAGGAG 360  
Db 441 GAGGTGGAGGTGGTGGGAAATTCGGGGGCGACGACAAAGCCCAAAATCAAGGAG 500  
Qy 361 GAGGTGGAGGTGGTGGGAAATTCGGGGGCGACGACAAAGCCCAAAATCAAGGAG 420  
Db 501 GAGGTGGAGGTGGTGGGAAATTCGGGGGCGACGACAAAGCCCAAAATCAAGGAG 560  
Qy 421 ATCTGTCAGCGGGCTTACCTGGGGTGTCTTCATCATCTTCGCGCGGCTGTGTGCTG 480  
Db 561 ATCTGTCAGCGGGCTTACCTGGGGTGTCTTCATCATCTTCGCGCGGCTGTGTGCTG 620  
Qy 481 TACACCGGCAAGATCTCATCGCTGCTGTGATGAGAGAAATGAAGACGGCGAGTGTG 540  
Db 621 TACACCGGCAAGATCTCATCGCTGCTGTGATGAGAGAAATGAAGACGGCGAGTGTG 680  
Qy 541 CGCGTGGGAGCTTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600  
Db 681 CGCGTGGGAGCTTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 740  
Qy 601 CTGGGCGGCGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660  
Db 741 CTGGGCGGCGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 800  
Qy 661 TACGTGGTGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 720  
Db 801 TACGTGGTGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 860  
Qy 721 AGTCTGTGTCATATTCGAGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780  
Db 861 AGTCTGTGTCATATTCGAGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 920  
Qy 781 AAGGCGGCTTCAAGTTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840  
Db 921 AAGGCGGCTTCAAGTTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 980  
Qy 841 GTATAGCTTACTGTCTATCGGGGCGGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAG 900  
Db 981 GTATAGCTTACTGTCTATCGGGGCGGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAG 1040  
Qy 901 ATCGAGCTCAAGAGTTCCTCATCTCCATTCGATTCGATTCGATTCGATTCGATTCGAT 960  
Db 1041 ATCGAGCTCAAGAGTTCCTCATCTCCATTCGATTCGATTCGATTCGATTCGATTCGAT 1100  
Qy 961 ATCTTCTGCTTCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020  
Db 1101 ATCTTCTGCTTCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1160  
Qy 1021 AATGAGCAGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
Db 1161 AATGAGCAGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1220  
Qy 1081 AATGAGCAGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140  
Db 1221 AATGAGCAGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1280  
Qy 1141 GTGGTCAACATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200  
Db 1281 GTGGTCAACATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1340  
Qy 1201 GCTGTGAGGTGCTGAGAGTGTCTTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260  
Db 1341 GCTGTGAGGTGCTGAGAGTGTCTTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1400

Qy 1261 TGCTACAGCGCGAGCGGCGCTTGAAGTCTCTGGGGCTGAGCGTCTGCGCTGCGCGCTCGTC 1320  
Db 1401 TGCTACAGCGCGAGCGGCGCTTGAAGTCTCTGGGGCTGAGCGTCTGCGCTGCGCGCTCGTC 1460  
Qy 1321 GTCTTACAGCTGCTCATGGCCATTTATGTGCGGCACTTTCGGGCTGCTCATGGGCTCAC 1380  
Db 1461 GTCTTACAGCTGCTCATGGCCATTTATGTGCGGCACTTTCGGGCTGCTCATGGGCTCAC 1520  
Qy 1381 GGCAGCTTACAGGCGCGGCGCTCTGTTCTTCTGCGCCAGCCTTTTCACCTTGGCGCTG 1440  
Db 1521 GGCAGCTTACAGGCGCGGCGCTCTGTTCTTCTGCGCCAGCCTTTTCACCTTGGCGCTG 1580  
Qy 1441 CTCTGGGCAAGCTGCTGTGGCACAAGTCTTCTTTCGACGTCGCGCATCTTCTGTCATCGGC 1500  
Db 1581 CTCTGGGCAAGCTGCTGTGGCACAAGTCTTCTTTCGACGTCGCGCATCTTCTGTCATCGGC 1640  
Qy 1501 GGCATCTGACGCTGCTGCGGCTTCTGTCATCTTTCGAGGGCTCATCGAAGCTTACCGA 1560  
Db 1641 GGCATCTGACGCTGCTGCGGCTTCTGTCATCTTTCGAGGGCTCATCGAAGCTTACCGA 1700  
Qy 1561 ACCAAGCGGAGGACTAG 1578  
Db 1701 ACCAAGCGGAGGACTAG 1718

RESULT 7  
AX259431  
LOCUS AX259431 1800 bp DNA linear PAT 26-OCT-2001  
DEFINITION Sequence 1 from Patent WO0173015.  
ACCESSION AX259431  
VERSION AX259431.1 GI:16508618  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
Rippmann, F. and Duecker, K.  
Identification of human gaba transporter  
Patent: WO 0173015-A 1 04-OCT-2001;  
MERCK PATENT GmbH (DE)  
FEATURES  
Location/Qualifiers  
1..1800  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
100..1677  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD10188.1"  
/db\_xref="GI:16508618"

CDS  
/translation="MATLLRSLKSLNATSVSNKSKMSPFARMGFQATDEAVGF  
AHCDDLDPEHROGLQMDILKAEFGPDEGAEPVEGDIHYQRGSGAPLPSPSGKQV  
GGGEGFGHDKPKITANEAGWNTNIAQMFVLGLPFIILHGGYGLGLFIIPAAVVC  
YTGKLLIACLYEENEDGVRVSYVAIANACCAAPRTTGGVYVVAQITELVMT  
ILYVVSGLNATYNSPFGIPVSKWSIIATALLPFCALKNLKAYSKSLCTLAHFV  
INILVIALSRARDWAEKFKPIDVKKFPISIGIIVFSYQIPLPSLEGNMQPS  
EFHMMNTHIAACVLKGFALVAYLTWADEKTEVITNLPGRIRAVNIFLVAALL  
SYPLFPFAVEVLEKSLFQEGSRAPFPACYSYSGDGRKSLWGLTLRCLVALVFTLLMAIYV  
PHFALLMGLTGLTGAIGLCPFLPSLFLRLMLRKLHMQVFDVAIFVIGGICSVSGF  
VHSLEGLIEAVRTNAED"

ORIGIN  
Query Match 99.9%; Score 1576.4; DB 6; Length 1800;  
Best Local Similarity 99.9%; Pred. No. 2.7e-229;  
Matches 1577; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 ATGGCCACTTCTCGCGAGCAAGCTGTCCACAGTGGCCACCGTCTCGTGTCCAAAGTCC 60  
Db 100 ATGGCCACTTCTCGCGAGCAAGCTGTCCAAAGTGGCCACCGTCTCGTGTCCAAAGTCC 159  
Qy 61 CAGGCCAAGATGAGCGCATGTTCGCCAGGATGGGTTTTTCAGCGCGGCCACCGATGAGGAG 120







Qy	1261	TGCTACAGCGCGCAGCGGGCGCTGAAAGTCTTGGGGGCTGACGCTGCGCTGGCGGCTCGTC	1320
Db	1435	TGCTACAGCGCGCAGCGGGCGCTGAAAGTCTTGGGGGCTGACGCTGCGCTGGCGGCTCGTC	1494
Qy	1321	GTCTTCACGCTGCTCATGGGCATTTATGTGCCGACATTCGCGCTGCTCATGGGCTCACC	1380
Db	1495	GTCTTCACGCTGCTCATGGGCATTTATGTGCCGACATTCGCGCTGCTCATGGGCTCACC	1554
Qy	1381	GGGAGCCTCAGGGCGGGGCTCTGTTTCTTGTGTCGCCAGCCTCTTTCACCTGGCGCTG	1440
Db	1555	GGCAGCCTCAGGGCGGGGCTCTGTTTCTTGTGTCGCCAGCCTCTTTCACCTGGCGCTG	1614
Qy	1441	CTCTGGCGCAAGCTGCTGTGGCACCAGTCTTCTTCGACCTGCGCATCTTTCGTCATCGGC	1500
Db	1615	CTCTGGCGCAAGCTGCTGTGGCACCAGTCTTCTTCGACCTGCGCATCTTTCGTCATCGGC	1674
Qy	1501	GGCATCTGCAGCGTGCCTGTCGCTCGCTCGAGGGCCTCATCGAGGCTACCGA	1560
Db	1675	GGCATCTGCAGCGTGCCTGTCGCTCGCTCGAGGGCCTCATCGAGGCTACCGA	1734
Qy	1561	ACCAACGCGGAGGACTAG	1578
Db	1735	ACCAACGCGGAGGACTAG	1752

RESULT 5	AK055051	2550 bp	linear	PRI 30-JAN-2004
LOCUS	AK055051			
DEFINITION	Homo sapiens CDNA FLJ30489 fis, clone BRAWH2000142, highly similar to Rattus norvegicus vesicular GABA transporter (VGAT) mRNA.			
ACCESSION	AK055051			
VERSION	AK055051.1	GI:16549699		
KEYWORDS	oligo capping; fis (full insert sequence).			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			

1 Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R., Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H., Sekine, M., Oyamashi, M., Nishi, T., Shibahara, T., Tanaka, T., Iishi, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y., Nagahori, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M., Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T., Kikkawa, E., Omura, Y., Abe, K., Kamiyama, K., Katsura, N., Sato, K., Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H., Hirakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M., Muraoka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S., Yeoide, M., Hottu, T., Kusano, J., Kanehori, K., Takahashi, Fujii, A., Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R., Takeuchi, K., Arita, M., Imose, N., Mueshino, K., Yuuki, H., Ohshima, A., Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T., Shiohata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, T., Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H., Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T., Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M., Tashiro, H., Tanigami, A., Fujwara, T., Ono, T., Yamada, K., Fujii, Y., Ozaki, K., Hirao, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N., Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T., Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsumura, K., Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togaishi, T., Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J., Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K., Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R., Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isozaki, T. and Sugano, S. Complete sequencing and characterization of 21,243 full-length human cDNAs

Nat. Genet. 36 (1), 40-45 (2004)  
14707039

2 Nishi, T., Nakagawa, S., Senoh, A., Mizuguchi, H., Inagaki, H.,

REFERENCE  
AUTHORS

TITLE

JOURNAL  
PUBMED  
REFERENCE  
AUTHORS

TITLE	JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Ota, T., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Wagatsuma, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahashi, K., Masuho, Y., Nagai, K. and Isogai, T.	NEDO human cDNA sequencing project	Unpublished	Isogai, T., Otsuki, T. and Sugiyama, T.	Direct Submission	Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 252-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)	NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'- and one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: RAB and HRI.

TITLE	JOURNAL	PUBLISHED	REFERENCE	AUTHOR
1. The Effect of Temperature on the Rate of Reaction of Hydrogen Peroxide with Potassium Iodide	Journal of Chemical Education	1965	1	Smith, J. D.
2. The Effect of Concentration on the Rate of Reaction of Hydrogen Peroxide with Potassium Iodide	Journal of Chemical Education	1965	2	Smith, J. D.
3. The Effect of Catalyst on the Rate of Reaction of Hydrogen Peroxide with Potassium Iodide	Journal of Chemical Education	1965	3	Smith, J. D.
4. The Effect of Surface Area on the Rate of Reaction of Hydrogen Peroxide with Potassium Iodide	Journal of Chemical Education	1965	4	Smith, J. D.
5. The Effect of pH on the Rate of Reaction of Hydrogen Peroxide with Potassium Iodide	Journal of Chemical Education	1965	5	Smith, J. D.

REMARK  
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgapbe-remail.nih.gov](mailto:cgapbe-remail.nih.gov)  
 Tissue Procurement: Life Technologies, Inc.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: National Institutes of Health Intramural  
 Sequencing Center (NISC),  
 Gaithersburg, Maryland;  
 Web site: <http://www.nisc.nih.gov/>  
 Contact: [nisc\\_mgc@hgrl.nih.gov](mailto:nisc_mgc@hgrl.nih.gov)  
 Blakesley, R.W., Bouffard G.G., Green, K., Brinkley, C., Brooks, S.,  
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,  
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,  
 Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,  
 McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,  
 Turgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,  
 Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAK Plate: 115 Row: c Column: 18  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: matched mRNA gi: 17999519.

FEATURES  
source

Location/Qualifiers  
 1..2519  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="MGC:61677 IMAGE:5180816"  
 /tissue\_type="Brain, Lung, Testis, adult, pooled whole"  
 /clone\_lib="NIH\_MGC\_115"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-SPORT6"  
 1..2519  
 /gene="VIAAT"  
 /note="synonym: VGAT"  
 /db\_xref="LocusID:140679"  
 175..1752  
 /gene="VIAAT"  
 /codon\_start=1  
 /product="vesicular inhibitory amino acid transporter"  
 /protein\_id="AAH53582.1"  
 /db\_xref="GI:31566392"  
 /db\_xref="LocusID:140679"  
 /translations="MATLRSLSNVATSVSNKSQRMGFMFARMFOAATDEAYGF  
 AHGDDLDFFHQGLQMDILKAEPEPCDEGAEPVEGDIHYQRGSGAPLPPSGSKDQV  
 GGGEGFGHDKPKITAEAGWNVNNAIQGMFVLGLPVAIIHGGYVGLFLIIFAAVVCC  
 YTGKILACLYEENEGEVVRVDSVVAIANACCAPRFTPLGRVNVNAQIIELVNWC  
 ILTVVSGNLYNYSFPLPVQSWSNIIATAVLLPCAFNLKNAKVSFSLCTLAHFV  
 INTLVAICLSRARDWAEKVFYIDVKFPISIGIIVFSYTSQIFPLEGNMQOPS  
 EYFCMNWTHIAACVLKGLFALVYLITWADETKEVITDNLPGSIRAVVNI FLVAKALL  
 SYPLPFAAVLEKSLFOEGSRAPFACVSGDRKLSWGLTURLCALVFTLIMALVY  
 PHFALLMLGLSLTGAGLCFLPLPSLPHLLRLWRKLWHQVFPDVAIFVIGGICSVSGF  
 VHSLEGLIEAYRNAED"

## ORIGIN

Query Match 100.0%; Score 1578; DB 9; Length 2519;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-229;  
 Matches 1578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 ATGGCCACCTTGCTCCGACGACGCTGTCCAAAGTGGCCACGTCGTGTCCAAAGATCC 60  
 175 ATGGCCACCTTGCTCCGACGACGCTGTCCAAAGTGGCCACGTCGTGTCCAAAGATCC 234  
 61 CAGGCCAAGATGAGCGCATGTCGCGAGATCGGTTTCAGGCGGCCGAGATGAGGAG 120  
 235 CAGGCCAAGATGAGCGCATGTCGCGAGATCGGTTTCAGGCGGCCGAGATGAGGAG 294  
 121 GCGGTGGGCTTCGCGATTCGCGACGACCTCGACTTTGAGCACCGCCAGGCGCTTCGAGATG 180  
 295 GCGGTGGGCTTCGCGATTCGCGACGACCTCGACTTTGAGCACCGCCAGGCGCTTCGAGATG 354

181 GACATCTCTGAAAGCCGAGGAGAGCCCTCGGGGACGAGGCGCTGAAGCGCCGCTCGAG 240  
 355 GACATCTCTGAAAGCCGAGGAGAGCCCTCGGGGACGAGGCGCTGAAGCGCCGCTCGAG 414  
 241 GGAGACATTCATTATCAGCGAGGAGCAGCGGAGCTCTCTGCGCGCTCCGCGCTCAAGGAC 300  
 415 GGAGACATTCATTATCAGCGAGGAGCAGCGGAGCTCTCTGCGCGCTCCGCGCTCAAGGAC 474  
 301 CAGGTGGAGGTGGTGGCGAATTCGGGGGCCACGACAAAGCCCAAAATCACGGCGTGGGAG 360  
 475 CAGGTGGAGGTGGTGGCGAATTCGGGGGCCACGACAAAGCCCAAAATCACGGCGTGGGAG 534  
 361 GCAGGTGGACCTGACCAACGCGATCCAGGCGATGTCGTGTGGGCTTACCTTAGCC 420  
 535 GCAGGTGGACCTGACCAACGCGATCCAGGCGATGTCGTGTGGGCTTACCTTAGCC 594  
 421 ATCTCTGACGCGCGCTTACCTGGGGTTGTTTCTCATCATCTTCGCGCGCTTGTGTGCTGC 480  
 595 ATCTCTGACGCGCGCTTACCTGGGGTTGTTTCTCATCATCTTCGCGCGCTTGTGTGCTGC 654  
 481 TACACCGGCAAGATCTCATTCGCGTCCCTGTACGAGGAGAAATGAAGCGCGAGGTGGT 540  
 655 TACACCGGCAAGATCTCATTCGCGTCCCTGTACGAGGAGAAATGAAGCGCGAGGTGGT 714  
 541 CCGGTGGGACTCGTACGTCGCGCATAGCCAAAGCTGTCGCGCGCGCTTCCCAAG 600  
 715 CCGGTGGGACTCGTACGTCGCGCATAGCCAAAGCTGTCGCGCGCGCTTCCCAAG 774  
 601 CTGGGCGCGGAGTGGTGAACGTAGCGCAGATCATCGAGCTGTGTGATGAGCTGCATCCTG 660  
 775 CTGGGCGCGGAGTGGTGAACGTAGCGCAGATCATCGAGCTGTGTGATGAGCTGCATCCTG 834  
 661 TAGCTGGTGGTGAAGCGCAACCTCATGTATCAACAGCTTCCCGGGGCTGCGCGCTGTCGAG 720  
 835 TAGCTGGTGGTGAAGCGCAACCTCATGTATCAACAGCTTCCCGGGGCTGCGCGCTGTCGAG 894  
 721 AGTCTCTGTTCATTATCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTTCCCTTAAGAACCTC 780  
 895 AGTCTCTGTTCATTATCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTTCCCTTAAGAACCTC 954  
 781 AAGCGCGCTGTCGAAGTTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
 955 AAGCGCGCTGTCGAAGTTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1014  
 841 GTCATAGCTTACTGTCTATCGCGGGCGCGAGCTGGGCTGGGAGAGGTCAAAGTTCTAC 900  
 1015 GTCATAGCTTACTGTCTATCGCGGGCGCGAGCTGGGCTGGGAGAGGTCAAAGTTCTAC 1074  
 901 ATCGAGCTCAAGAGTTCCTCATCTCCATTTGGCATCATCGTGTTCAGCTACAGCTCTCAG 960  
 1075 ATCGAGCTCAAGAGTTCCTCATCTCCATTTGGCATCATCGTGTTCAGCTACAGCTCTCAG 1134  
 961 ATCTCTCTGCTTCTGCTGGAGGCAATATGACAGCGCCAGCGAGTTCACATGCATGATG 1020  
 1135 ATCTCTCTGCTTCTGCTGGAGGCAATATGACAGCGCCAGCGAGTTCACATGCATGATG 1194  
 1021 AACTGGACGACATCGACGCTCGGTGCTCAAGGGGCTCTTCGCGCTCGCTCGCTACCTC 1080  
 1195 AACTGGACGACATCGACGCTCGGTGCTCAAGGGGCTCTTCGCGCTCGCTCGCTACCTC 1254  
 1081 ACTGGGGCGACGAGACCAAGAGGTTCATACGGATAAAGCTGCGCGGCTCCATCCGGGC 1140  
 1255 ACTGGGGCGACGAGACCAAGAGGTTCATACGGATAAAGCTGCGCGGCTCCATCCGGGC 1314  
 1141 GTGGTCAACATCTTTCTGCTGGCAGGCGCTGTTGCTCTATCTCTCTGCGCATTTCTTGGC 1200  
 1315 GTGGTCAACATCTTTCTGCTGGCAGGCGCTGTTGCTCTATCTCTCTGCGCATTTCTTGGC 1374  
 1201 GCTGTGAGGTGCTGGAGAGTCTCTTTCCAGGAGGAGCGCGGCTTTTCCCGGC 1260  
 1375 GCTGTGAGGTGCTGGAGAGTCTCTTTCCAGGAGGAGCGCGGCTTTTCCCGGC 1434

469	Db	 CAGGCCAAGATGAGCGGCATGTTTCGCGCAGGATGGGTGTTTCAGCGCGCCACGGATGAGGAG	528
121	Qy	 GCGGTGGGCTTCGCGCATTTGCGACGACCTTCGACTTTTGAGCAGCGCCAGCGGCTGCAGATG	180
529	Db	 GCGGTGGGCTTCGCGCATTTGCGACGACCTTCGACTTTTGAGCAGCGCCAGCGGCTGCAGATG	588
181	Qy	 GACATCTCTAAAGCCGAGGGAGAGCCCTCGCGGGACGAGGGGCGCTGAAGCGCCGCTCGAG	240
589	Db	 GACATCTGAAAGCCGAGGGAGAGCCCTCGCGGGACGAGGGGCGCTGAAGCGCCGCTCGAG	648
241	Qy	 GGAGACATCCATTATCAGCGAGGACGCGGAGCTTCCTTCGCGCCCTTCGGGCTTCCAAGGAC	300
649	Db	 GGAGACATCCATTATCAGCGAGGACGCGGAGCTTCCTTCGCGCCCTTCGGGCTTCCAAGGAC	708
301	Qy	 CAGGTGGGAGTGGTGGCGAATTCGCGGGGCCAGCAAGCCCAAAATCACGGGCGTGGGAG	360
709	Db	 CAGGTGGGAGTGGTGGCGAATTCGCGGGGCCAGCAAGCCCAAAATCACGGGCGTGGGAG	768
361	Qy	 GCAGGCTGGAAAGTGACCAACGCCATCCAGGGGCATGTTGCTGTGGGCCCTACCTACGCC	420
769	Db	 GCAGGCTGGAAAGTGACCAACGCCATCCAGGGGCATGTTGCTGTGGGCCCTACCTACGCC	828
421	Qy	 ATCCTGCA CGGCGGCTACCTGGGGTGTGTTCTCATCATCTTCCGCGCGTGTGTGCTGCTC	480
829	Db	 ATCCTGCA CGGCGGCTACCTGGGGTGTGTTCTCATCATCTTCCGCGCGTGTGTGCTGCTC	888
481	Qy	 TACACGGCAAGATCCTCATCGCTGCTGTACGAGGAGATCAAGACGCGCAGGTGGTG	540
889	Db	 TACACGGCAAGATCCTCATCGCTGCTGTACGAGGAGATCAAGACGCGCAGGTGGTG	948
541	Qy	 CGGCTGGGACCTCGTACGTGGCCATAGCCAAAGCCCTGCTCGCCCGCGGCTTCCCAACG	600
949	Db	 CGGCTGGGACCTCGTACGTGGCCATAGCCAAAGCCCTGCTCGCCCGCGGCTTCCCAACG	1008
601	Qy	 CTGGGCGCCGAGTGGTGAACGTAGCGCAGATCATCGAGCTGTGTATGAGTGCATCTCTG	660
1009	Db	 CTGGGCGCCGAGTGGTGAACGTAGCGCAGATCATCGAGCTGTGTATGAGTGCATCTCTG	1068
661	Qy	 TAGTGGTGTGAGTGGCAACCTCATGTACACAGCTTCCCGGGCTGCCCGTGCAG	720
1069	Db	 TAGTGGTGTGAGTGGCAACCTCATGTACACAGCTTCCCGGGCTGCCCGTGCAG	1128
721	Qy	 AAGTCTGTGTCATATTTCGCCAGCGCGTGTCTGTGCTTGCSCCTTCTTTAAGAACCTC	780
1129	Db	 AAGTCTGTGTCATATTTCGCCAGCGCGTGTCTGTGCTTGCSCCTTCTTTAAGAACCTC	1188
781	Qy	 AAGCGCGTGTCCAAGTTTCAGTGTGTGCACTCTTGCGCCACTTGTGTCATCAATATCTCTG	840
1189	Db	 AAGCGCGTGTCCAAGTTTCAGTGTGTGCACTCTTGCGCCACTTGTGTCATCAATATCTCTG	1248
841	Qy	 GTCAATAGCTACTGTCTATCGGGGCGCGACTGGGCGCTGGGAGAGGTCAAGTTCTAC	900
1249	Db	 GTCAATAGCTACTGTCTATCGGGGCGCGACTGGGCGCTGGGAGAGGTCAAGTTCTAC	1308
901	Qy	 ATCGACGTCAAGAAATTCCTCAATGGGCATCATCGTGTTCAGCTACACGTCTCAG	960
1309	Db	 ATCGACGTCAAGAAATTCCTCAATGGGCATCATCGTGTTCAGCTACACGTCTCAG	1368
961	Qy	 ATCTTCTGTGCTTCGCTGGAGGGCAATATGACGAGCCGACGAGTTTCCAATGATG	1020
1369	Db	 ATCTTCTGTGCTTCGCTGGAGGGCAATATGACGAGCCGACGAGTTTCCAATGATG	1428
1021	Qy	 AATGGAGCAGCATCGCAGCTGCTGCTCAAGGGCCTCTTCCGGCTCGTGCCTACCTC	1080
1429	Db	 AATGGAGCAGCATCGCAGCTGCTGCTCAAGGGCCTCTTCCGGCTCGTGCCTACCTC	1488
1081	Qy	 ACCTGGGCGGACGAGACCAAGAGGTTCATACGGATAACTGCGCGGCTCCATCCGCGCC	1140
1489	Db	 ACCTGGGCGGACGAGACCAAGAGGTTCATACGGATAACTGCGCGGCTCCATCCGCGCC	1548
1141	Qy	 GTGCTCAACATCTTCTGTGGGCCAAGGCGCTGTGTCTATCCTCTGCAATCTTTGCTC	1200

Query Match									
Best Local Similarity 100.0%; Score 1578; DB 9; Length 1592;									
Matches 1578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	ATGGCCACTTGTCTCCGAGCAAGCTGTCTCAACGTGGCCACGTGCTGTCACAAAGTCC	60						
DB	6	ATGGCCACTTGTCTCCGAGCAAGCTGTCTCAACGTGGCCACGTGCTGTCACAAAGTCC	65						
QY	61	CAGGCCAAGATGAGCGGCATGTCGCCAGGATGGTTTTTCAGGGCGGCCACCGATGAGG	120						
DB	66	CAGGCCAAGATGAGCGGCATGTCGCCAGGATGGTTTTTCAGGGCGGCCACCGATGAGG	125						
QY	121	CGGTGGGCTTTCGGGCATTCGCGACGACCTTCGACTTTGAGCACCGCCAGGGCTCGACATG	180						
DB	126	CGGTGGGCTTTCGGGCATTCGCGACGACCTTCGACTTTGAGCACCGCCAGGGCTCGACATG	185						
QY	181	GACATCTGAAAGCCGAGGGAGAGCCCTGCGGGACGAGGCGCTGGAAGCGCCCGTCCAG	240						
DB	186	GACATCTGAAAGCCGAGGGAGAGCCCTGCGGGACGAGGCGCTGGAAGCGCCCGTCCAG	245						
QY	241	GGAGACATCCATTATCAGCGAGGCGAGGAGCTCTCTGCGCCCTTCGGCTCCAGGAC	300						
DB	246	GGAGACATCCATTATCAGCGAGGCGAGGAGCTCTCTGCGCCCTTCGGCTCCAGGAC	305						
QY	301	CAGGTGGAGGTGTGGCGAATTCGGGGGCCACGACAAAGCCCAAAATCACGCGCTGGAG	360						
DB	306	CAGGTGGAGGTGTGGCGAATTCGGGGGCCACGACAAAGCCCAAAATCACGCGCTGGAG	365						
QY	361	GCAGGTGGAAAGCTGACCAAGCCATCAGGGCATGTTGCTGCTGGGCTTACCCCTACGCC	420						
DB	366	GCAGGTGGAAAGCTGACCAAGCCATCAGGGCATGTTGCTGCTGGGCTTACCCCTACGCC	425						
QY	421	ATCTGTGACGCGCTACCTGGGCTGTCTTCATCATCTTCGCGCGCTGTGTGTCTGC	480						
DB	426	ATCTGTGACGCGCTACCTGGGCTGTCTTCATCATCTTCGCGCGCTGTGTGTGTCTGC	485						
QY	481	TACACCGGCAAGATCTCTCATCGCTGCTGTACGAGGAGATGAAGACGGCGAGTGGTG	540						
DB	486	TACACCGGCAAGATCTCTCATCGCTGCTGTACGAGGAGATGAAGACGGCGAGTGGTG	545						
QY	541	CGGTGGGGAATCGTACGTGGCCATAGCCAAAGCCCTGCTGCGCCCGCGCTTCCCAAGC	600						
DB	546	CGGTGGGGAATCGTACGTGGCCATAGCCAAAGCCCTGCTGCGCCCGCGCTTCCCAAGC	605						
QY	601	CTGGGCGCCGAGTGGTGAAGTGAAGCAGATCATCGAGCTGTGTGATGAGTGCATCTG	660						
DB	606	CTGGGCGCCGAGTGGTGAAGTGAAGCAGATCATCGAGCTGTGTGATGAGTGCATCTG	665						
QY	661	TACGTGTGTGTGAGTGGCACTCATGTACAAAGCTTCCGCGGCTGCGCGTGTGTCAG	720						
DB	666	TACGTGTGTGTGAGTGGCACTCATGTACAAAGCTTCCGCGGCTGCGCGTGTGTCAG	725						
QY	721	AAGTCTGTGTCATTATCGCCACGCGCGTGTGCTGCTTGGCGCTTCTTAAAGAACCTC	780						
DB	726	AAGTCTGTGTCATTATCGCCACGCGCGTGTGCTGCTTGGCGCTTCTTAAAGAACCTC	785						
QY	781	AAGCCGTGTGCAAGTTCAGTCTGCTGCACTTCGCGCCACCTTCGTCATCAATATCTG	840						
DB	786	AAGCCGTGTGCAAGTTCAGTCTGCTGCACTTCGCGCCACCTTCGTCATCAATATCTG	845						
QY	841	GTCAATAGCTTACTGTCTATCGCGGCGCGGACCTGGGCGCTGGAGAGGTCAAGTTCCTAC	900						
DB	846	GTCAATAGCTTACTGTCTATCGCGGCGCGGACCTGGGCGCTGGAGAGGTCAAGTTCCTAC	905						
QY	901	ATCGAGCTCAAGAGTTCCTCCATCTCCATTTGGCATCATCGTGTTCAGTTCAGTCTTCAG	960						
DB	906	ATCGAGCTCAAGAGTTCCTCCATCTCCATTTGGCATCATCGTGTTCAGTTCAGTCTTCAG	965						
QY	961	ATCTTCTGCTGCTGCTGGAGGGCAATATGAGGAGCCGACGAGTTCCTCATGATGATG	1020						
DB	966	ATCTTCTGCTGCTGCTGGAGGGCAATATGAGGAGCCGACGAGTTCCTCATGATGATG	1025						
QY	1021	AACTGGACGCACATCGCAGCTGCTGCTCAAGGGCTCTTTCGGCTCGTTCGCTACCTC	1080						
Db	1026	AACTGGACGCACATCGCAGCTGCTGCTCAAGGGCTCTTTCGGCTCGTTCGCTACCTC	1085						
QY	1081	ACCTGGGCGGACGAGACCAAGAGGTTCATACGATAACTTCGCCGCTCCATCCGCGCC	1140						
DB	1086	ACCTGGGCGGACGAGACCAAGAGGTTCATACGATAACTTCGCCGCTCCATCCGCGCC	1145						
QY	1141	GTGGTCAACATCTTCTGTGGGCAAGGCGCTGTGTCTTCTATCTCTGCCATTTCTTGCC	1200						
DB	1146	GTGGTCAACATCTTCTGTGGGCAAGGCGCTGTGTCTTCTATCTCTGCCATTTCTTGCC	1205						
QY	1201	GCTGTCCAGGTGTGGAGAAAGTCCTTCCAGGAAGCGAGCGCGCTTTTTCGCCGCC	1260						
DB	1206	GCTGTCCAGGTGTGGAGAAAGTCCTTCCAGGAAGCGAGCGCGCTTTTTCGCCGCC	1265						
QY	1261	TGCTACAGCGGCGACGCGCGCTGAAGTCTCTGGGGGTGACGTGCGCTGCGCGCTCGTC	1320						
DB	1266	TGCTACAGCGGCGACGCGCGCTGAAGTCTCTGGGGGTGACGTGCGCTGCGCGCTCGTC	1325						
QY	1321	GTCTTACGCTGTCTCATGGCCATTTATGTGCGGCACTTCGCGCTGCTCATGGGCTCAC	1380						
DB	1326	GTCTTACGCTGTCTCATGGCCATTTATGTGCGGCACTTCGCGCTGCTCATGGGCTCAC	1385						
QY	1381	GGAGCGCTCACGCGCGCGCGCTCTGTTCTTGTGCGCCAGCTCTTTCACCTGCGCGCTG	1440						
DB	1386	GGAGCGCTCACGCGCGCGCGCTCTGTTCTTGTGCGCCAGCTCTTTCACCTGCGCGCTG	1445						
QY	1441	CTCTGGCGCAAGCTGTGTGGCACCAAGTCTTCTTGCAGCTGCGCATCTTTCGTCATCGGC	1500						
DB	1446	CTCTGGCGCAAGCTGTGTGGCACCAAGTCTTCTTGCAGCTGCGCATCTTTCGTCATCGGC	1505						
QY	1501	GGCATCTGACGCTGTCCGCGCTTCTGTCACCTCCCTCGAGGCGCTCATCGAAGCGCTACCGA	1560						
DB	1506	GGCATCTGACGCTGTCCGCGCTTCTGTCACCTCCCTCGAGGCGCTCATCGAAGCGCTACCGA	1565						
QY	1561	ACCAACGCGGAGGACTAG 1578							
DB	1566	ACCAACGCGGAGGACTAG 1583							
RESULT 3									
AX398635									
LOCUS									
AX398635 2261 bp DNA linear PAT 27-MAY-2002									
DEFINITION									
Sequence 3 from Patent WO0220753.									
ACCESSION									
AX398635									
VERSION									
AX398635.1 GI:21261223									
KEYWORDS									
SOURCE									
Homo sapiens (human)									
ORGANISM									
Homo sapiens									
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
REFERENCE									
1									
AUTHORS									
Fridde, C.J., Gerhardt, B. and Hu, Y.									
TITLE									
Human gaba transporter protein and polynucleotides encoding the									
same									
JOURNAL									
Patent: WO 0220753-A 3 14-MAR-2002;									
Lexicon Genetics Incorporated (US)									
FEATURES									
source									
1..2261									
/organism="Homo sapiens"									
/mol_type="unassigned DNA"									
/db_xref="taxon:9606"									
ORIGIN									
Query Match									
Best Local Similarity 100.0%; Score 1578; DB 6; Length 2261;									
Matches 1578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	ATGGCCACTTGTCTCCGAGCAAGCTGTCTCAACGTGGCCACGTGCTGTCACAAAGTCC	60						
DB	409	ATGGCCACTTGTCTCCGAGCAAGCTGTCTCAACGTGGCCACGTGCTGTCACAAAGTCC	468						
QY	61	CAGGCCAAGATGAGCGGCATGTCGCCAGGATGGTTTTTCAGGGCGGCCACCGATGAGG	120						



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 3, 2004, 18:32:15 ; Search time 6606 Seconds  
(without alignments)  
11296.268 Million cell updates/sec

Title: US-09-940-919-1  
Perfect score: 1578  
Sequence: 1 atggccacctgtctcgagcag.....gaaccaacgaggagactag 1578

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues  
Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl: \*  
1: gb\_ba: \*  
2: gb\_btg: \*  
3: gb\_in: \*  
4: gb\_on: \*  
5: gb\_ov: \*  
6: gb\_pat: \*  
7: gb\_ph: \*  
8: gb\_pl: \*  
9: gb\_pr: \*  
10: gb\_ro: \*  
11: gb\_sts: \*  
12: gb\_sy: \*  
13: gb\_un: \*  
14: gb\_vi: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1578	100.0	1578	6	AX398633 Sequence
2	1578	100.0	1592	9	AX044836 Homo sapi
3	1578	100.0	2261	6	AX398635 Sequence
4	1578	100.0	2519	9	BC053582 Homo sapi
5	1578	100.0	2550	9	AK055051 Homo sapi
6	1576.4	99.9	1763	6	AX497262 Sequence
7	1576.4	99.9	1800	6	AX259431 Sequence
8	1574.8	99.8	1734	6	AX497264 Sequence
9	1574.8	99.8	1734	6	AX497268 Sequence
10	1574.8	99.8	1734	6	AX497270 Sequence
11	1574.8	99.8	1734	6	AX497272 Sequence
12	1567.4	99.3	1733	6	AX497266 Sequence
13	1518.8	96.2	2756	9	AB062931 Macaca fa
14	1424.8	90.3	2174	9	BC036458 Homo sapi
15	1358	86.1	1473	6	CQ727095 Sequence
16	1352.2	85.7	2392	6	AR438018 Sequence
17	1352.2	85.7	2392	10	AF030253 Rattus no
18	1345.8	85.3	2814	10	BC052020 Mus muscu
19	1344.2	85.2	1648	6	AX147507 Sequence

20	1323.4	83.9	1989	9	AK091079	AK091079 Homo sapi
21	1315.4	83.4	1889	10	MMVIAAT	AJ001598 Mus muscu
22	1193.2	75.6	96299	9	AL133519	AL133519 Human DNA
c 23	1140	72.2	1723	6	AX147533	AX147533 Sequence
24	1087	68.9	85427	2	HSDJ470K3	AL109655 Homo sapi
25	1043.8	66.1	6668	10	AB080232	AB080232 Mus muscu
26	1043.8	66.1	196078	10	AL663091	AL663091 Mouse DNA
c 27	1043.8	66.1	247690	2	AC095186	AC095186 Rattus no
28	1043.8	66.1	249669	2	AC137204	AC137204 Rattus no
29	929	58.9	230452	2	AC136118	AC136118 Rattus no
30	908.6	57.6	2751	5	BC075429	BC075429 Xenopus t
31	881	55.8	2949	5	BC057733	BC057733 Xenopus t
32	476.4	30.2	494	9	AF489843	AF489843 Macaca mu
33	366	23.2	1650	6	CQ590550	CQ590550 Sequence
34	366	23.2	3722	6	CQ590549	CQ590549 Sequence
35	366	23.2	9625	2	AC017264	AC017264 Drosophil
c 36	366	23.2	165177	3	AC007588	AC007588 Drosophil
c 37	366	23.2	229684	3	AE003815	AE003815 Drosophil
38	360	22.8	360	11	HS470K3T	AB030975 H.sapiens
39	296	18.8	2028	3	AB158401	AB158401 Ciona int
40	238.2	15.1	819	5	EX931551	EX931551 Gallus ga
41	232.2	14.7	2740	3	AK115620	AK115620 Ciona int
42	137.2	8.7	1586	3	AF031935	AF031935 Caenorhab
43	137.2	8.7	1586	6	AR438017	AR438017 Sequence
c 44	120.4	7.6	38199	3	CBRNW14H7	AC084685 Caenorhab
c 45	112	7.1	48095	3	CET20G5	Z304233 Caenorhabdi

ALIGNMENTS

RESULT 1  
AX398633  
LOCUS AX398633 1578 bp DNA linear PAT 27-MAY-2002  
DEFINITION Sequence 1 from Patent WO0220753.  
ACCESSION AX398633  
VERSION AX398633.1 GI:21261222  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1  
AUTHORS Friddle,C.J., Gerhardt,B. and Hu,Y.  
TITLE Human gaba transporter protein and polynucleotides encoding the same  
JOURNAL Patent: WO 0220753-A 1 14-MAR-2002;  
Lexicon Genetics Incorporated (US)  
FEATURES  
source  
1..1578  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 1578; DB 6; Length 1578;  
Best Local Similarity 100.0%; Pred. No. 1.6e-229;  
Matches 1578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGCCACCTTGTCTCCGAGCAGCTGTCCAACTGGCCACGTCCTGTGTCCAACTGTC 60  
Db 1 ATGGCCACCTTGTCTCCGAGCAGCTGTCCAACTGGCCACGTCCTGTGTCCAACTGTC 60  
QY 61 CAGGCCAAGATGAGCGGCATGTTCCAGGATGGGTTTTTCAGCGGCCACGATGAGGAG 120  
Db 61 CAGGCCAAGATGAGCGGCATGTTCCAGGATGGGTTTTTCAGCGGCCACGATGAGGAG 120  
QY 121 GCGGTGGGTTTCCGCGATTTCGACGACCTTCGACTTTGAGCACCGCCAGGSCCTGCAGATG 180  
Db 121 GCGGTGGGTTTCCGCGATTTCGACGACCTTCGACTTTGAGCACCGCCAGGSCCTGCAGATG 180  
QY 181 GACATCCTGAAGCCAGGAGGAGCCCTCGCGGGAGCGAGCCCTCGAGGCGCTGAAGCGCCGCTCGAG 240